UNCERTAINTY QUANTIFICATION AND INTEGRATION IN ENGINEERING SYSTEMS

By

Shankar Sankararaman

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Civil Engineering

May, 2012

Nashville, Tennessee

Approved:

Professor Sankaran Mahadevan

Professor Prodyot K. Basu

Professor Bruce Cooil

Professor Gautam Biswas

Professor Mark P. McDonald

To my Parents and Teachers

ACKNOWLEDGMENTS

First and foremost, I express my heartfelt gratitude to my advisor Prof. Sankaran Mahadevan, for his expert professional advice and guidance throughout my graduate studies at Vanderbilt University. His constant encouragement has been a great source of motivation not only in my academic career but also in my personal life. I am extremely indebted to him, for all that I have learnt during my graduate education, and look forward to continuing this relationship for the rest of my professional career.

I would also like to acknowledge the help and support I received from my committee members Prof. Prodyot K. Basu, Prof. Gautam Biswas, Prof. Bruce Cooil, and Prof. Mark P. McDonald, who provided useful comments and suggestions during this research. I express my sincere thanks to the various sponsor agencies that funded this research and my mentors Dr. Lee D. Peterson and Dr. Samuel C. Bradford (NASA Jet Propulsion Laboratory), Dr. Kai Goebel (NASA Ames Research Center), Dr. Angel Urbina (Sandia National Laboratory), Prof. Jayathi Murthy (University of Texas, Austin), Prof. Yongming Liu (Clarkson University), Mark Derriso (Air Force Research Laboratory), and Lawrence Green (NASA Langley Research Center) for their advice within various research projects that effectively contributed to this dissertation. I am also grateful for the teaching assistantship from the Department of Civil and Environmental Engineering at Vanderbilt University during my first year in graduate school. Further, the computational resources at Vanderbilt University's ACCRE (Advanced Computing Center for Research and Education) were valuable for the simulations conducted in this research.

I would like to thank all the students and faculty members of the Reliability and Risk Engineering and Management doctoral program at Vanderbilt University, with

whom I was engaged in many technical discussions that not only helped this dissertation but also led to individual professional development. I am grateful to the following members of my research group for assistance in several technical accomplishments: You Ling (fatigue crack growth analysis, implementation of Bayesian inference and sampling techniques), Kyle McLemore (implementation of the test resource allocation methodology), Dr. Chris Shantz (generation of the finite element model for fatigue crack growth analysis), and Chen Liang (model development for the dynamics challenge problem). I also thank Dr. Sirisha Rangavajhala, Dr. Vadiraj Hombal, Dr. John McFarland, and Dr. Sohini Sarkar, with whom I have had several valuable technical discussions. I also acknowledge the wonderful company, co-operation, and encouragement of my colleagues and friends, Dr. Venkata Sura and Dr. Kais Zaman, during my stay at Vanderbilt.

This dissertation would not have been possible but for the boundless support of my family. I thank my father S. Sankararaman, mother R. Anuradha, and my wife Poornima Murthi for their constant encouragement and exemplary patience during my graduate education. Poornima also directly helped me with writing this dissertation, by proofreading a few chapters, improving the quality of several figures, and formatting and typesetting the overall document. Last but not the least, I thank my mathematics teacher, Late Prof. V. Balakrishnan (retired Principal at Tagore Arts College, Pondicherry, India), whose exemplary teaching methods have equipped me with a profound understanding of the fundamental concepts in Mathematics.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

1.1 Motivation

Research in the area of uncertainty quantification and the application of stochastic methods to the study of engineering systems has gained considerable attention during the past twenty years. This can be attributed to the necessity and desire to design engineering systems with increasingly complex architectures and new materials. These systems can be multi-level, multi-scale, and multi-disciplinary in nature, and may need to be decomposed into simpler components and subsystems to facilitate efficient model development, analysis and design. The development and implementation of computational models is not only sophisticated and expensive, but also based on physics which is often not well-understood. The study of engineering systems is further complicated by several issues: (1) presence of various sources of error and uncertainty that need to be rigorously accounted for; (2) limited availability of full-scale system tests/data; (3) budget and time constraints that limit modeling and/or testing resources; and (4) availability of information at multiple levels, that need to be carefully and meaningfully integrated for uncertainty quantification and decision-making purposes.

A comprehensive framework for the treatment of uncertainty is essential to facilitate design, risk assessment and management, inspection and maintenance scheduling, and operational decision-making in engineering systems. Quantification of margins and uncertainties (QMU) has been advocated as a systematic framework to aid system-level uncertainty quantification and decision making under uncertainty [1].

This dissertation focuses on advancing the state of the art in uncertainty quantification methods, in order to facilitate quantification of margins and uncertainty, and aid risk-informed decision-making in engineering systems. Several computational methods are proposed to overcome different challenges in the areas of uncertainty representation, quantification, and integration in hierarchical system models. A Bayesian approach is pursued for the treatment of uncertainty; this approach emphasizes the use of subjective probabilities rather than objective probabilities. While the concept of objective probabilities is based on the relative frequency of occurrence of an event, the concept of subjective probabilities is based on the analyst's degree of belief regarding the event.

A major advantage of the Bayesian approach is that it provides an efficient computational tool in the form of a Bayesian network (sometimes, simply referred to as Bayes network or Bayes net), which facilitates the integration of multiple computational models, various sources of uncertainty and errors, and experimental data, towards overall uncertainty quantification of the system-level prediction. This dissertation makes use of the Bayesian approach and proposes new computational methods for quantification of the various types of uncertainty, integration of the various sources of uncertainty across multiple models and thereby, provides information in order to facilitate risk-informed decision-making during the different stages of the life cycle of engineering systems.

The rest of this chapter develops the different objectives of this dissertation. Sections 1.2 and 1.3 briefly discuss the various aspects of uncertainty quantification and integration respectively. Section 1.4 delineates the various research objectives and Section 1.5 enumerates the significant contributions and highlights of this research. Section 1.6 describes the organization of this dissertation.

1.2 Uncertainty Quantification

The sources of uncertainty in the design, analysis, and operation of engineering systems are broadly classified into two types. The first type, aleatory uncertainty, is due to the presence of physical variability and inherent randomness in nature. If the outcome of an experiment differs each time the experiment is run, then this is an example of aleatory uncertainty. This type of uncertainty is irreducible. The second type, epistemic uncertainty, is due to lack of knowledge regarding a particular quantity and/or a physical phenomenon. This type of uncertainty could be reduced (and sometimes eliminated) if and when new information is available. While the topic of aleatory uncertainty has been extensively studied for the past fifty years, the topic of epistemic uncertainty has been gaining significant attention only during the recent years. There are several types of situations where epistemic uncertainty may be present, and correspondingly, researchers have pursued different types of mathematical approaches (both probabilistic and non-probabilistic) for the treatment of epistemic uncertainty.

One important objective of this dissertation is to develop a unified framework for the representation and quantification of uncertainty. This dissertation makes a more detailed classification of the various types of uncertainty in engineering systems, as follows:

1.2.1 Physical Variability

As mentioned earlier, this type of uncertainty is referred to as aleatory uncertainty. The inputs to the engineering system may be random (for e.g. fluctuations in loading), the parameters may have variability (for e.g. elastic modulus), and this leads to an uncertain output. It is common to represent such random variables using probability distributions.

1.2.2 Data Uncertainty

Data uncertainty is of two major types. The most commonly considered type of data uncertainty is measurement errors (both at the input and output levels). It may be argued that measurement errors occur due to natural variability and hence, must be classified as a type of aleatory uncertainty. Since measurement errors are associated with data collection, this dissertation classifies measurement error as a type of data uncertainty; however, the treatment of measurement errors is similar to that of aleatory uncertainty. The second type of data uncertainty occurs during the characterization of variability. Sometimes, the available data (sparse point data and/or intervals) may be insufficient to precisely estimate the distribution parameters of probability distributions, thereby leading to statistical uncertainty. This is an example of epistemic uncertainty (i.e. uncertainty reducible in the light of new information); if sufficient data is available, then the distribution parameters can be estimated precisely.

1.2.3 Model Uncertainty

The engineering system under study is represented using a mathematical model, and the corresponding mathematical equations are solved using computer codes. Model uncertainty is an example of epistemic uncertainty and comprises of three different types of errors/uncertainty. First, the intended mathematical equation is solved using a computer code which leads to rounding off errors, solution approximation errors and coding errors; this issue is addressed in the model verification process. Second, the model parameters may not be readily known, and field data may need to be collected in order to calibrate them; this issue is addressed in the model calibration process. Third, the mathematical equation itself may not be an accurate representative of reality, which leads to model form error; this issue is addressed in the model

validation process. It is essential to account for the sources of physical variability and data uncertainty during model verification, validation, and calibration activities. The combined effect of solution approximation errors, model form errors, and model parameter uncertainty is referred to as the overall model uncertainty.

1.2.4 Goals in Uncertainty Quantification

The overall goal in uncertainty quantification is to mathematically represent and quantify the various sources of uncertainty (both aleatory and epistemic) and compute the combined effect of the various types of uncertainty on the system-level response. The various activities related to uncertainty quantification are (1) identification and quantification of the different types and sources of uncertainty; (2) model verification, validation, and calibration leading to quantification of associated model errors and uncertainty; and (3) quantifying the combined effect of the various sources of uncertainty using an uncertainty propagation method. Note that model verification can be performed prior to testing; on the other hand, test data is essential for both model validation and calibration.

1.3 Uncertainty Integration

Sometimes, engineering systems are developed hierarchically, based on component, subsystem, and system-level models. Hence, several computational models are developed, and these models, in turn, form the basis for the analysis and design of the system as a whole. The various sources of uncertainty discussed in Section 1.2 are present in each level of the system hierarchy, and hence, the various uncertainty quantification activities such as verification, validation, calibration, uncertainty propagation, etc. need to be performed at each level of the hierarchy. The system-level

prediction is dependent on the information available at all lower levels, and hence, it is important to integrate the results from the uncertainty quantification activities at multiple levels, in order to quantify the overall uncertainty in the system-level response prediction.

1.3.1 Hierarchical System Configurations

The development of the methodology for the integration of uncertainty quantification activities depends upon the interaction between the various models in the system hierarchy. These models may combine and/or interact in several possible ways, leading to three fundamentally different configurations. Note that a typical engineering system may consist of a combination of such configurations; however if the fundamental methodology for the treatment of each of these configurations is developed, then it is possible to generalize such methodology to any complex combination of configurations.

1. **Non-sequential Hierarchy:** In this type of hierarchy, a set of parameters are common across multiple levels of modeling; there is no direct relation between the outputs of these models. Note that each level may have its own set of "local" parameters and/or inputs in addition to the common parameters. Typically, along the hierarchy, the complexity of the underlying physical phenomenon increases, and consequentially, the complexity of the computational model and the cost of testing increase. For example, a coupon, a beam, and a plate made of the same material constitute a non-sequential hierarchy.

2. **Sequential Hierarchy or Feed-forward Coupling:** In this type of hierarchy, the output of a lower-level model becomes an input to a higher-level model. Usually, the two levels consist of different physics, thereby rendering the system

multi-disciplinary. For example, the rise in the temperature of a wire due to heat conduction leads to a change in the resistance and hence, the current carrying capacity of the wire.

3. **Feedback Coupling:** In this type of hierarchy, the two models (usually, governing two different physics) are interconnected in such a way that the input of one becomes an output to the other; hence, from a hierarchical point of view, both the models are considered to be "at the same level". For example, in fluid-structure analysis, the displacement (output of structural analysis) and pressure (output of fluid analysis) fields are dependent on each other. Feedback coupling necessitates iterative analysis between the two models until both the solutions converge.

Feedback coupling and feed-forward coupling are also referred to as strong and weak coupling respectively, in the literature [2].

1.3.2 Goals in Uncertainty Integration

In an engineering system, information is available in the form of input data, output data, models, expert opinion, etc. at multiple levels. The goal in uncertainty integration is to efficiently integrate all the above information in an uncertainty quantification framework, and provide information in order to guide in system analysis, design, testing operations, and maintenance scheduling.

1.4 Research Objectives

The overall goal of this dissertation is to develop a framework for uncertainty quantification and integration, in order to provide information necessary for risk-informed decision-making during various stages in the life-cycle of engineering systems. The

various objectives of this dissertation are grouped under two headings: uncertainty quantification and uncertainty integration.

1.4.1 Uncertainty Quantification

The first major objective is to develop a rigorous framework for uncertainty quantification and propagation with respect to component-level or subsystem-level analysis which require single-level models. Statistical methods are developed for the treatment of data uncertainty and model uncertainty. Then, the effects of the various sources of uncertainty are combined in uncertainty propagation to quantify the uncertainty in the output. Uncertainty quantification must also be accompanied by a sensitivity analysis of the various sources of uncertainty. The method of global sensitivity analysis has been used by several researchers to quantify the sensitivity of output to the inputs and parameters; this dissertation refers to this analysis as forward sensitivity analysis. The proposed methods for data uncertainty quantification and model uncertainty quantification are applied to fatigue crack growth analysis under uncertainty, as a case study. The various activities related to uncertainty quantification are addressed under the following topics:

1. Data Uncertainty due to sparse and interval data
2. Model Uncertainty - Model verification, validation, and calibration
3. Case Study - Fatigue crack growth analysis

1.4.2 Uncertainty Integration

The second major objective focuses on the integration of uncertainty quantification activities across the different hierarchies of modeling discussed earlier in Section 1.3.1. A Bayesian network-based approach is developed to integrate multiple

models, various sources of uncertainty including model errors, and testing data, in order to quantify the uncertainty in the system-level response. One challenge with respect to the use of a Bayesian network is that it cannot account for feedback coupling due to its acyclic nature. Hence, a new decoupled approach is developed for multi-disciplinary analysis, in order to address this challenge. If the tests available at lower levels can be used to quantify the uncertainty in the system-level prediction, then it should be possible to prioritize tests in order to achieve a reasonable level of reduction in the variance of the system-level prediction. Hence, the Bayesian network approach is used to guide resource allocation for test selection and prioritization. It is observed that some parameters are more sensitive to the data; further, this issue is complicated by the presence of other types of uncertainty. Hence, a rigorous framework is developed to study the sensitivity of the calibration parameter to the various sources of uncertainty and the data; this task is referred to as "inverse sensitivity analysis" in this dissertation. The various activities related to integration of uncertainty are addressed under the following topics:

1. Integration of uncertainty quantification activities
2. Multi-disciplinary systems analysis
3. Resource allocation for test prioritization
4. Inverse sensitivity analysis

1.5 Highlights of the Dissertation: What's New?

1. A new likelihood-based methodology is developed for the representation and quantification of epistemic uncertainty due to sparse and interval data (Chapter III). Three cases – (1) single distribution type (Section 3.4); (2) multiple

competing distribution types (Section 3.6); and (3) non-parametric type without any assumption regarding distribution type (Section 3.7) – are considered.

2. A global sensitivity analysis-based methodology is developed in order to quantify the individual contributions of variability and epistemic uncertainty (distribution type and parameter uncertainty). The methodology is developed for quantifying individual contributions in a single variable as well as contributions to the output of a response function (Sections 3.5, 3.6.1.3 and 3.6.2.4).

3. A statistical methodology is developed for model calibration under uncertainty, by including different types of uncertainty and different types of data situations such as uncharacterized data, imprecise data, etc. (Section 4.3.8).

4. Two methods – Bayesian hypothesis testing (Sections 4.4.1) and model reliability method (Section 4.4.2) – are investigated for model validation under uncertainty, and extended to different scenarios, including interval data, uncharacterized data, time series data, etc.

5. The methods for the quantification of data uncertainty and model uncertainty are applied to fatigue crack growth analysis (Chapter V), as a case study. Previous work in probabilistic fracture mechanics has considered only physical variability, and some sources of model uncertainty. Chapter V discusses parameter estimation, model validation, and crack growth prediction by considering different sources of uncertainty. These methods are developed for crack growth in structures with complicated geometry and multi-axial variable amplitude loading.

6. A "roll-up" methodology is developed to integrate the results of model verification, validation, and calibration activities in order to quantify the uncertainty

in the system-level prediction (Chapter VI). This methodology is developed for sequential and non-sequential hierarchical configurations.

7. Multi-disciplinary systems analysis with feedback coupling is considered in Chapter VII, where a new likelihood-based methodology for uncertainty quantification and propagation. A decoupled approach is developed to facilitate the application of Bayesian networks to systems with feedback coupling.

8. A Bayesian network-based optimization methodology is developed for cost-effective prioritization of tests in multi-level systems, in order to meet the variance reduction target in system-level prediction (Chapter VIII).

9. A new "inverse sensitivity analysis" methodology is developed to analyze the sensitivity of model parameters to the other sources of uncertainty and data (Chapter IX). This is significantly different from the existing methods of global sensitivity analysis which analyze the sensitivity of model output to model inputs and parameters. This inverse sensitivity analysis methodology is useful to identify important contributors of uncertainty in model calibration. This is useful for dimension reduction and design of experiments. Section 9.4 develops the methodology for a single-level model, and Section 9.5 extends the methodology to hierarchical systems built using multiple models.

1.6 Organization of the Dissertation

The remainder of this dissertation is organized into several chapters where computational methods are developed in order to address the research objectives in Section 1.4. Prior to the development of new methods, Chapter II provides a review of well-known topics in the area of uncertainty quantification. These topics include a formal introduction to the fundamental probability theory, the multiple interpretations

of probability, elementary concepts in the Bayesian methodology, Bayesian networks, uncertainty propagation and reliability analysis, global sensitivity analysis, Markov Chain Monte Carlo sampling for statistical inference, and Gaussian process surrogate modeling. Original research contributions are presented through Chapters III – IX.

Chapter III discusses the treatment of data uncertainty, due to the presence of sparse point data and interval data for uncertainty quantification including representation and propagation. Data uncertainty due to measurement errors are discussed in the subsequent chapter. This is because measurement errors are related to data from testing; such data is used to quantify model form uncertainty, which is the focus of Chapter IV. Further, the method of forward global sensitivity analysis is used to analyze the contributions of the various sources of uncertainty to the overall uncertainty in the model prediction. The methods proposed for uncertainty quantification are applied to fatigue crack growth analysis in Chapter V, as a case study. Chapters III and IX address the issue of uncertainty quantification and propagation, and consider only single-level models.

Systems represented using multiple models are considered in Chapters VI – VIII, where the focus is on uncertainty integration. Chapter VI focuses on two types of hierarchical system models, namely sequential and non-sequential. A Bayesian network-based computational methodology is developed in order to quantify the system-level prediction uncertainty, by integrating various uncertainty quantification activities including model verification, validation, and calibration. The Bayesian network is acyclic and cannot explicitly include feedback coupling in multi-disciplinary analysis. Therefore, Chapter VII proposes a new likelihood-based decoupled methodology for uncertainty quantification in multi-disciplinary systems, using which the feedback coupling can be replaced with feed-forward coupling, and thus included in the Bayesian network. Chapter VIII develops an optimization-based methodology for

prioritization of testing, when component and subsystem tests are used to reduce the uncertainty in the system-level response.

Testing aids in the calibration of parameters and thereby in the reduction of uncertainty. The uncertainty in the calibration parameters is affected by the presence of other sources of uncertainty in the system. In Chapter IX, a new inverse sensitivity analysis methodology is developed to analyze the sensitivity of model calibration under uncertainty. The methodology is first developed for a single-level model and then extended to systems with multiple models. Finally, Chapter X presents a few summarizing remarks and suggests recommendations for future work.

Each of the above chapters review the relevant literature and discuss the current state of art prior to the discussion of the proposed methods. In each individual chapter, the notations and symbols used for method development are introduced and explained. The numerical examples presented in various chapters illustrate the proposed methods using engineering applications in civil, mechanical, and aerospace structures.

CHAPTER II

BACKGROUND

2.1 Overview

This chapter reviews a few background concepts, which are fundamental to the development of original research contributions of this dissertation. The goal is not to provide a detailed tutorial on these topics, but only to discuss the basic governing mathematical equations, and the relevant references.

The various topics in this chapter are organized as follows. Section 2.2 formally introduces the fundamental theory of probability, and Section 2.3 discusses the two different interpretations (frequentist and subjectivist) of probability. While classical statistics is based on the frequentist interpretation of probability, the subjectivist interpretation is the backbone of Bayesian statistics. As stated earlier in Chapter I, this dissertation uses the Bayesian paradigm for uncertainty quantification and integration, and therefore Section 2.4 provides a brief introduction to the Bayesian methodology. The concept of a Bayesian network is introduced and it is explained how the Bayesian network can be used for solving forward problems (uncertainty propagation) and inverse problems (statistical inference). Computational methods for uncertainty propagation are discussed in Section 2.5; sampling-based methods and analytical methods are both considered to quantify the output uncertainty by propagating the uncertainty in the inputs through a computational model. Section 2.6 reviews the global sensitivity analysis methodology which is recommended to be performed in tandem with uncertainty propagation; the focus in global sensitivity analysis is to apportion the output uncertainty to the uncertainty in the inputs. The

solution of statistical inverse problems is conventionally facilitated through the use of Markov Chain Monte Carlo sampling (since analytical methods are computationally infeasible), which is presented in Section 2.7. Finally, it is explained that the above methods for uncertainty propagation and statistical inference require several thousands of evaluations of the computational models, and therefore, it may be necessary to build surrogate models to replace them. Section 2.8 discusses one specific type of surrogate model, i.e. Gaussian process surrogate model, which will be used to replace expensive computational models in the rest of this dissertation.

2.2 Fundamentals of Probability Theory

The fundamental theory of probability is well-established in the literature, including many textbooks and journal articles. The roots of probability lie in the analysis of games of chance by Gerolamo Cardano in the sixteenth century, and by Pierre de Fermat and Blaise Pascal in the seventeenth century.

In earlier days, researchers were interested only in discrete probabilities, and with the advent of mathematical analysis, the importance of continuous probabilities steadily increased. This led to a significant change in the understanding and formal definition of probability. The classical definition of probability was based on counting the number of favorable outcomes, and it was understood that this definition cannot be applied to continuous probabilities (refer to Bertrand's Paradox [3]). Hence, the modern definition of probability, which is based on set theory and functional mapping, is more commonly used in recent times.

Discrete probability deals with events where the sample space is countable. Consider the sample space (Ω), which is equal to the set of all possible outcomes. The modern definition of probability maps every element $x \in \Omega$ to a "probability value" $f(x)$ such that:

1. $f(x) \in [0,1] \; \forall \; x \in \Omega$

2. $\sum\limits_{x \in \Omega} f(x) = 1$

Any event E can be expressed as a subset of the sample space Ω ($E \in \Omega$), and the probability of the event E is defined as:

$$P(E) = \sum_{x \in E} f(x) \tag{2.1}$$

Hence, the function $f(x)$ is a mapping from a point in the sample space to a probability value, and is referred to as probability mass function (PMF).

Continuous probability theory deals with cases where the sample space is continuous and hence uncountable; consider the case where the set of outcomes of a random experiment is equal to the set of real numbers ($\mathbb{R}$). In this case, the modern definition of probability is in terms of the cumulative distribution function (CDF), defined as $F_X(x) = P(X \leq x)$, i.e. the CDF of the random variable X evaluated at x is equal to the probability that the random variable X can take on a value less than or equal to x. This CDF necessarily follows the following properties:

1. $F_X(x)$ is monotonically non-decreasing, and right continuous.

2. $\underset{x \to -\infty}{Lim} \; F_X(x) = 0$

3. $\underset{x \to \infty}{Lim} \; F_X(x) = 1$

If the function $F_X(x)$ is is absolutely continuous and differentiable, then the derivative of the CDF is denoted as the probability density function (PDF) $f_X(x)$. Therefore,

$$f_X(x) = \frac{dF_X(x)}{dx} \tag{2.2}$$

For any set $E \subseteq \mathbb{R}$, the probability of the random variable X being in E can be written as:

$$P(X \in E) = \int_{x \in E} dF_x(x) \tag{2.3}$$

If the PDF exists, then

$$P(X \in E) = \int_{x \in E} f_X(x)dx \tag{2.4}$$

Note that the PDF exists only for continuous random variables, whereas the CDF exists for all random variables (including discrete variables) whose realizations belong to $\mathbb{R}$. A PDF or CDF is said to be valid if and only if it satisfies all of the above properties. The above discussion can be easily be extended to multiple dimensions by considering the space $\mathbb{R}^n$.

The above principles of probability theory are not only fundamental to this dissertation, but will be repeatedly used in several chapters. For instance, in Chapter III, the focus is on constructing probability distributions, and hence all of the above equations will be used (sometimes as optimization constraints, as in Eq. 3.30).

2.3 Interpretations of Probability

The previous section formally defined probability in terms of cumulative distribution function and probability density function. What is the meaning of this probability? Though the concepts of probability are well-established in the literature, there is considerable disagreement among researchers on the interpretation of probability. There are two major interpretations based on physical and subjective probabilities respectively. It is essential to understand the difference between these two interpretations before delving deeper into this dissertation; this is mainly because the latter philosophy is widely used in this research work. In fact, this dissertation advocates the

latter philosophy because of its ability to integrate the various sources of uncertainty across multiple levels of models that represent the overall engineering system.

2.3.1 Physical Probability

Physical probabilities [4], also referred to objective or frequentist probabilities, are related to random physical systems such as rolling dice, tossing coins, roulette wheels, etc. Each trial of the experiment leads to an event (which is a subset of the sample space), and in the long run of repeated trials, each event tends to occur at a persistent rate, and this rate is referred to as the "relative frequency". These relative frequencies are expressed and explained in terms of physical probabilities. Thus, physical probabilities are defined *only* in the context of random experiments. The theory of classical statistics is based on physical probabilities. Within the realm of physical probabilities, there are two types of interpretations: von Mises' frequentist [5] and Popper's propensity [6]; the former is more easily understood and widely used.

In the context of physical probabilities, the mean of a random variable, sometimes referred to as the population mean, is deterministic. It is meaningless to talk about the PDF of this mean. In fact, for any type of parameter estimation, the underlying parameter is assumed to be deterministic and only an estimate of this parameter is obtained. The uncertainty in the parameter estimate is addressed through confidence intervals. The interpretation of confidence intervals is, at times, confusing and misleading, and the uncertainty in the parameter estimate cannot be used for further uncertainty quantification. For example, if the uncertainty in the elastic modulus was estimated using a simple axial test, this uncertainty cannot be used for quantifying the response in a plate made of the same material. This is a serious limitation, since it is not possible to propagate uncertainty after parameter estimation, which is often necessary in the case of model-based quantification of uncertainty in the system-level

response. Another disadvantage of this approach is that, when a quantity is not random, but unknown, then the tools of probability cannot be used to represent this type of uncertainty (epistemic). The second interpretation of probability, i.e. the subjective interpretation, overcomes these limitations.

2.3.2 Subjective Probability

Subjective probabilities [7] can be assigned to any "statement". It is not necessary that the concerned statement is in regard to an event which is a possible outcome of a random experiment. In fact, subjective probabilities can be assigned even in the absence of random experiments. The Bayesian methodology is based on subjective probabilities, which are simply considered to be degrees of belief and quantify the extent to which the "statement" is supported by existing knowledge and available evidence. Calvetti and Somersalo [8] explain that "randomness" in the context of physical probabilities is equivalent to "lack of information" in the context of subjective probabilities.

In this approach, even deterministic quantities can be represented using probability distributions which reflect the subjective degree of the analyst's belief regarding such quantities. As a result, probability distributions can be assigned to parameters that need to be estimated, and therefore, this interpretation facilitates uncertainty propagation after parameter estimation; this is helpful for uncertainty integration across multiple models.

For example, consider the case where a variable is assumed to be normally distributed and it is desired to estimate the mean and the standard deviation based on available point data. If sufficient data were available, then it is possible to uniquely estimate these distribution parameters. However, in some cases, data may be sparse and therefore, it may be necessary to quantify the uncertainty in these distribution

parameters. Note that this uncertainty is an example of epistemic uncertainty; the quantities may be estimated deterministically with enough data. The former philosophy based on physical probabilities inherently assumes that these distribution parameters are deterministic and expresses the uncertainty through confidence intervals on mean and standard deviation. It is not possible to propagate this description of uncertainty through a mathematical model. On the other hand, the Bayesian methodology can calculate probability distributions for the distribution parameters, which can be easily used in uncertainty propagation. Therefore, the Bayesian methodology provides a framework in which epistemic uncertainty can be also addressed using probability theory, in contrast with the frequentist approach.

The fundamentals of Bayesian philosophy are well-established in several textbooks [9–12], and the Bayesian approach is being increasingly applied to engineering problems in recent times, especially to solve statistical inverse problems. In this dissertation, the Bayesian methodology is extensively used to integrate not only the different types and sources of uncertainty, but also to integrate multiple models which represent the overall system under study.

2.4 The Bayesian Methodology

This section introduces the Bayesian methodology, since this approach will be extensively used in the remainder of the dissertation. The fundamental concepts alone are explained here; advanced concepts will be explained as and when necessary in later chapters.

2.4.1 Bayes Theorem

Though named after the 18^{th} century mathematician and theologian Thomas Bayes [13], it was the French mathematician Pierre-Simon Laplace who pioneered and popularized what is now called Bayesian probability [14, 15]. For a brief history of Bayesian methods, refer to [16]. The law of conditional probability is fundamental to the development of Bayesian philosophy:

$$P(AB) = P(A|B)P(B) = P(B|A)P(A) \tag{2.5}$$

Consider a list of mutually exclusive and exhaustive events A_i ($i = 1$ to n) that together form the sample space. Let B denote any other event from the sample space such that $P(B) > 0$. Based on Eq. 2.5, it follows that:

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_j P(B|A_j)P(A_j)} \tag{2.6}$$

What does Eq. 2.6 mean? Suppose that the probabilities of events A_i ($i = 1$ to n) are known to be equal to $P(A_i)$ ($i = 1$ to n) before conducting any random experiments. These probabilities are referred to as prior probabilities in the Bayesian context. Suppose that a random experiment has been conducted and event B has been observed. In the light of this data, the so-called posterior probabilities $P(A_i|B)$ ($i = 1$ to n) can be calculated using Eq. 2.6.

The quantity $P(B|A_i)$ is the probability of observing the data conditioned on A_i. It can be argued that, event B has "actually been observed", and there is no uncertainty regarding its occurrence, which renders the probability $P(B|A_i)$ meaningless. Hence, researchers "invented" new terminology in order to denote this quantity. In earlier days, this quantity was referred to as "inverse probability", and since the

advent of Fisher [17, 18] and Edwards [19], this terminology has become obsolete, and has been replaced by the term "likelihood". In fact, it is also common to write $P(B|A_i)$ as $L(A_i)$.

2.4.2 Bayesian Inference

The concept of Bayes theorem can be extended from the discrete case to the continuous case. Consider the context of statistical inference where a set of parameters $\boldsymbol{\theta}$ needs to be inferred. All the current knowledge regarding this parameter is represented in the form of a prior distribution denoted by $f'(\boldsymbol{\theta})$. The choice of the prior distribution reflects the subjective knowledge of uncertainty regarding the variable before any observation. It is assumed that the prior distribution is able to explain the data with some degree of uncertainty; in other words, there exists a non-empty set E such that $\forall\ \theta \in E$, the prior PDF and likelihood values evaluated $\forall\ \theta \in E$ are both non-zero.

Measurement data (D) is collected on a quantity which depends on the parameter ($\boldsymbol{\theta}$). This information is then used to update the distribution of θ to arrive at the posterior distribution ($f''(\boldsymbol{\theta})$), as:

$$f''(\boldsymbol{\theta}) = \frac{L(\boldsymbol{\theta})f'(\boldsymbol{\theta})}{\int L(\boldsymbol{\theta})f'(\boldsymbol{\theta})d\boldsymbol{\theta}} \tag{2.7}$$

In Eq. 2.7, $L(\boldsymbol{\theta})$ is the likelihood function of $\boldsymbol{\theta}$ and is proportional to $P(D|\boldsymbol{\theta})$, i.e. probability of observing the data D conditioned on the parameter $\boldsymbol{\theta}$. The concept of likelihood will be briefly explained in the following subsection.

The denominator on the RHS of Eq. 2.7 is simply a normalizing constant, which ensures that $f''(\boldsymbol{\theta})$ is a valid PDF. So, Eq. 2.7 is sometimes written as:

$$f''(\boldsymbol{\theta}) \propto L(\boldsymbol{\theta})f'(\boldsymbol{\theta}) \tag{2.8}$$

The posterior in Bayesian inference is always known only up to a proportionality constant and it is necessary generate samples from this posterior for uncertainty analysis. When there is only one parameter, the proportionality constant can be calculated through one-dimensional integration. Often, multiple parameters may be present, and hence, multi-dimensional integration may not be affordable to calculate the proportionality constant. Therefore, a class of methods popularly referred to as Markov Chain Monte Carlo (MCMC) sampling is used to generate samples from the Bayesian posterior. In general, these methods can be used when it is desired to generate samples from a PDF which is known only up to a proportionality constant. The topic of MCMC will be discussed in detail later in this chapter, in Section 2.7.

2.4.3 Notes on the Likelihood Function

The likelihood function is defined as the probability of observing data conditioned on the parameters, i.e. $L(\boldsymbol{\theta}) = P(D|\boldsymbol{\theta})$; note that, since the *data* (D) *has actually been observed*, the terminology "probability of observing the data" is physically meaningless. Therefore, as explained earlier in Section 2.4.1, this quantity was renamed as "the likelihood". The likelihood function does not follow the laws of probability discussed in Section 2.2, and must not be confounded with probability distributions or distribution functions. In fact, Edwards [19] explains that the likelihood function is meaningful only up to a proportionality constant; the relative values of the likelihood function are alone significant and the absolute values are not of interest.

The concept of likelihood is used both in the context of physical probabilities (frequentist) and subjective probabilities, especially in the context of parameter estimation. In fact, Edwards [19] refers to the likelihood method as the third or middle way.

From a frequentist point of view (the underlying parameters are deterministic),

the likelihood function can be maximized in order to obtain the maximum likelihood estimate of the parameters. According to Fisher [20], the popular least squares approach is parameter estimation is an indirect approach to parameter estimation and one can "solve the real problem directly" by maximizing the "probability of observing the given data" conditioned on the parameter θ [20, 21]. Further, it is also possible to construct likelihood-based confidence intervals for the inferred parameters [22].

On the other hand, the likelihood function can also be interpreted using subjective probabilities. Singpurwalla [23, 24] explains that the likelihood function can be viewed as a collection of weights or masses and therefore, is meaningful only up to a proportionality constant [19]. In other words, if $L(\theta^{(1)}) = 10$, and $L(\theta^{(2)}) = 100$, then it is 10 ten times more likely for $\theta^{(2)}$ than $\theta^{(1)}$ to correspond to the observed data. The entire likelihood function can be used in Bayesian inference, as in Eq. 2.7, in order to obtain the entire PDF of the parameters.

2.4.4 Bayesian Network

An attractive feature of the Bayesian approach is the ability to integrate the various sources of uncertainty in a systematic manner, through a Bayesian network (also referred to as Bayes network or simply Bayes net). Further, the Bayesian network also allows the integration of the multiple models, associated sources of uncertainty, and experimental data. A Bayesian network [25, 26] is a graphical representation of conditional probability relationships between uncertain quantities. Each uncertain quantity is represented as a node and successive nodes are connected to each other using unidirectional arrows that express dependence in terms of conditional probabilities. If the probability distribution of one uncertain variable Z depends on the realization of another uncertain variable Y, then the nodes corresponding to the variables Y and Z are said to be parent and child nodes with respect to each other. While

Bayesian networks are constructed hierarchically, the node corresponding to Z may become a parent to one or more child nodes.

Fig. 2.1 shows a conceptual Bayesian network that aids in uncertainty quantification across multiple levels of models, and observed data. Circles correspond to uncertain variables and squares represent observed data. A solid line arrow represents a conditional probability link, and a dashed line arrow represents the link of a variable to its observed data if available.

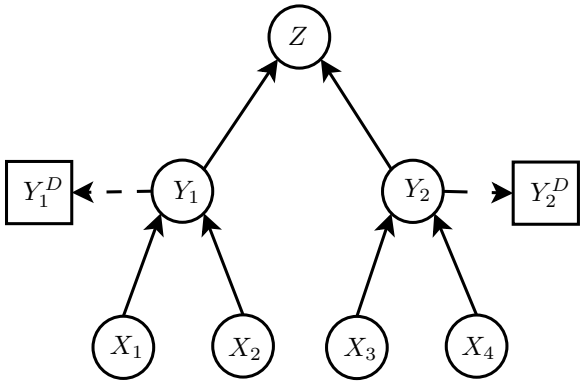


Figure 2.1: Bayesian Network Illustration

In Fig. 2.1, a system level output Z is a function of two subsystem level quantities Y_1 and Y_2; in turn, Y_1 is a function of subsystem-level input X_1 and model parameter X_2, and similarly, Y_2 is a function of subsystem-level input X_3 and model parameter X_4. For example, in a beam deflection study, the applied force is an input, the elastic modulus is a model parameter, while the deflection is measured and a model is built to predict the same. Experimental data Y_1^D and Y_2^D are available for comparison with the respective model predictions Y_1 and Y_2.

The Bayesian network is useful in solving both forward problems (uncertainty propagation along the flow of the Bayesian network) and inverse problems (updating the uncertainty in the parent nodes based on the data on a dependent quantity (child node), using Bayesian inference).

2.4.4.1 Uncertainty Propagation: Forward Problem

In the forward problem, the probability distributions of the inputs (X_1 and X_3 in Fig. 2.1) and model parameters (X_2 and X_4 in Fig. 2.1) are known or assumed, and these distributions are used to calculate the probability density function (PDF) of Y_1 and Y_2, which in turn are used to calculate the PDF of the system-level output Z as:

$$\begin{aligned} f_Z(z) &= \int f_Z(z|y_1, y_2) f_{Y_1}(y_1) f_{Y_2}(y_2) dy_1 dy_2 \\ f_{Y_1}(y_1) &= \int f_{Y_1}(y_1|x_1, x_2) f_{X_1}(x_1) f_{X_2}(x_2) dx_1 dx_2 \\ f_{Y_2}(y_2) &= \int f_{Y_2}(y_2|x_3, x_4) f_{X_3}(x_3) f_{X_4}(x_4) dx_3 dx_4 \end{aligned} \tag{2.9}$$

Eq. 2.9 can be solved using forward uncertainty propagation methods, explained later in Section 2.5.

2.4.4.2 Inference: Inverse Problem

In the inverse problem, the probability densities of the model parameters (X_2 and X_4 in Fig. 2.1) can be updated based on the observed data (Y_1^D and Y_2^D) using Bayes theorem as:

$$f_{X_2,X_4}(x_2, x_4|Y_1^D, Y_2^D) \propto L(x_2, x_4) f_{X_2}(x_2) f_{X_4}(x_4) \tag{2.10}$$

In Eq. 2.10, the prior distributions of the model parameters X_2 and X_4 are given by $f_{X_2}(x_2)$ and $f_{X_4}(x_4)$ respectively. The joint posterior density is given by $f_{X_2,X_4}(x_2, x_4|Y_1, Y_2)$. The likelihood function $L(x_2, x_4)$ is calculated as the probability of observing the given data (Y_1^D, Y_2^D), conditioned on the parameters being updated, i.e. x_2 and x_4. The likelihood function $L(\boldsymbol{\theta})$ needs to explicitly account for the uncertainty in the inputs X_1 and X_3; this aspect is explained in detail later in Chapters IV and IX.

Thus, the Bayesian network can be used for both forward problems (estimating the PDF of Z) and inverse problems (calibrating parameters X_2 and X_4 based on data).

Section 2.5 and 2.7 describe methods for uncertainty propagation (forward problem) and statistical inference (inverse problem) respectively. While sampling methods and analytical methods are discussed in the former case, the latter case considers only Markov Chain Monte Carlo-based sampling methods.

2.5 Methods for Uncertainty Propagation

Consider a generic computational model $Y = G(\boldsymbol{X})$, which is used to represent the performance of an engineering system. The input is a vector and hence denoted in bold as $\boldsymbol{X}$, whereas the output Y is a scalar. The model G is deterministic, i.e. for a given realization of $\boldsymbol{X}$, there is a corresponding output, which is a realization of Y. The inputs $\boldsymbol{X}$ are uncertain, and this leads to uncertainty in the output Y. A generic realization of $\boldsymbol{X}$ is denoted as $\boldsymbol{x}$, and a generic realization of Y is denoted as y.

The goal in uncertainty propagation is to propagate the input uncertainty through G, in order to the calculate the entire CDF $F_Y(y)$, and the corresponding PDF $f_Y(y)$. The entire CDF can be calculated as:

$$F_Y(y) = \int_{g(\boldsymbol{X})<y} f_{\boldsymbol{X}}(\boldsymbol{x}) d\boldsymbol{x} \tag{2.11}$$

It is harder to write a similar expression for PDF calculation, although the following equation attempts to.

$$f_Y(y) = \int f_Y(y|\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}) d\boldsymbol{x} \tag{2.12}$$

In Eq. 2.12, the domain of integration is such that $f_{\boldsymbol{X}}(\boldsymbol{x}) \neq 0$. In fact, this practice is employed in the entire dissertation, i.e. whenever the integrand contains a PDF, then the domain of integration is equal to the domain of the PDF. Note that Eq. 2.12 is not

very meaningful because y is single-valued given $\boldsymbol{x}$, and hence $f_Y(y|\boldsymbol{x})$ is nothing but a Dirac delta function. Alternatively, the PDF can be calculated by differentiating the CDF, as:

$$f_Y(y) = \frac{dF_Y(y)}{dy} \tag{2.13}$$

Two types of methods – sampling-based and analytical methods – are be used to calculate the PDF and CDF of Y.

2.5.1 Monte Carlo Sampling

The most intuitive method for uncertainty propagation is to make use of Monte Carlo simulation (MCS). In this method, several random realizations of $\boldsymbol{X}$ are generated based on CDF inversion, and the corresponding random realizations of Y are computed. Then the CDF $F_Y(y)$ is calculated as the proportion of the number of realizations where the output realization is less than y. The generation of each realization requires one evaluation/simulation of the computational model. Several thousands of realizations may often be needed to calculate the entire CDF, especially for very high/low values of y. Error estimates for the CDF, in terms of the number of simulations, are available in the literature [27, 28]. Once the samples of Y are obtained, then a histogram can be drawn easily or the empirical PDF of Y can be calculated using kernel density estimation [29].

Advanced MCS methods such as importance sampling, stratified sampling, latin hypercube sampling, etc. are also available to aid in the reduction of computational effort [28, 30]. The basic underlying concept of these methods is to generate pseudo-random numbers which are uniformly distributed on the interval [0, 1]; then the CDF of $\boldsymbol{X}$ is inverted to generate the corresponding realization of $\boldsymbol{X}$. Therefore, these methods are applicable only when the CDF of X is fully known and can be inverted. This is the case in uncertainty propagation because the PDFs of $\boldsymbol{X}$ are assumed to

be known and the goal is to propagate them through the model $Y = G(\boldsymbol{X})$. The topic of uncertainty propagation will be considered in all of the forthcoming chapters in this dissertation, and the above sampling methods will be repeatedly used for this purpose, and kernel density estimation will be used to construct the PDF based on Monte Carlo samples.

There is another class of sampling methods, collectively referred to as MCMC sampling methods, which are used to draw samples from a probability distribution whose CDF cannot be inverted or whose PDF is known only up to a proportionality constant. This is often used to solve statistical inverse problems (Bayesian inference), and hence discussed later in Section 2.7.

2.5.2 Analytical Methods

A new class of methods was developed by reliability engineers in order to facilitate efficient, quick but approximate calculation of the CDF $F_Y(y)$; the focus is not on the calculation of the entire CDF function but only to evaluate the CDF at a particular value of the output, i.e. $F_Y(Y = y_c)$; the value of y_c is chosen so that the CDF value, i.e. the probability $P(Y \leq y_c)$ is the failure probability of the system represented by the model $Y = G(\boldsymbol{X})$.

The basic concept is to "linearize" the model G so that the the output Y can be expressed as a linear combination of the random variables. Further, the random variables are transformed into uncorrelated standard normal space and hence, the output Y is also a normal variable (since the linear combination of normal variables is normal). Therefore, the CDF value $F_Y(Y = y_c)$ can be computed using the standard normal distribution function. The transformation of random variables $\boldsymbol{X}$ into uncorrelated standard normal space ($\boldsymbol{U}$) is denoted by $\boldsymbol{U} = T(\boldsymbol{X})$, and the details of the transformation can be found in [28].

Since the model G is non-linear, the failure probability depends on the location of "linearization". This linearization is done at the so-called most probable point (MPP) which is the shortest distance from origin to the limit state, calculated in the $\boldsymbol{U}$—space. Then, the failure probability is calculated as $P_f = \Phi(-\beta)$, where Φ denotes the standard normal CDF function, and β denotes the aforementioned shortest distance. The MPP and the shortest distance are estimated using the well-known Rackwitz-Fiessler algorithm [31], which is based on a repeated linear approximation of the non-linear constraint $G(\boldsymbol{x}) - y_c = 0$. This method is popularly known as the first-order reliability method (FORM). There are also several second order reliability methods (SORM) based on the quadratic approximation of the limit state [28, 32–34].

The entire CDF can be calculated using repeated FORM analyses by considering different values of y_c; for example, if FORM is performed at 10 different values of y_c, the corresponding CDF values are calculated, and an interpolation scheme can be used to calculate the entire CDF, which can be differentiated to obtain the PDF. This approach is difficult because it is almost impossible to choose such multiple values of y_c, because the range (i.e. extent of uncertainty) of Y is unknown. This difficulty is overcome by the use of an inverse FORM method [35] where multiple CDF values are chosen and the corresponding values of y_c are calculated. This approach is simpler because it is easier to choose multiple CDF values since the range of CDF is known to be [0, 1].

In this dissertation, FORM and inverse FORM are not implemented for the purpose of uncertainty propagation. However, Chapter VII will require the use of FORM for the numerical implementation of a new methodology for uncertainty quantification in multi-disciplinary analysis. Hence, details regarding the Rackwitz-Fiessler algorithm (identification of MPP and calculation of $F_Y(y_c)$) will be provided in Chapter VII.

2.6 Global Sensitivity Analysis

The previous section discussed methods for propagating the uncertainty in the inputs X through a computational model $Y = G(\boldsymbol{X})$, in order to compute the uncertainty in the model output Y. Note that G is a deterministic transfer function from a given realization of $\boldsymbol{X}$ to a realization of Y. The topic of sensitivity analysis is closely associated with uncertainty propagation, and Saltelli et al. [36] state that uncertainty propagation is incomplete without the results of quantitative sensitivity analysis. The goal in sensitivity analysis is to apportion the uncertainty in Y to the uncertainty in inputs $\boldsymbol{X}$.

An intuitive approach to sensitivity analysis would be to ignore the uncertainty in each input quantity, one at a time, i.e. replace the probability distribution of the input quantity by a deterministic value and calculate the reduction in variance in the model output. This approach would only give the local sensitivity and the result would depend on the chosen deterministic value. Saltelli et al. [36] explain that local sensitivities are not sufficient to study the contributions of multiple sources of uncertainty to the overall prediction uncertainty and it is necessary to pursue a global sensitivity analysis approach for this purpose. The term "global" refers to computing the sensitivity metric considering the entire probability distribution of the input.

The fundamental theorem that governs the development of the global sensitivity analysis methodology is the variance decomposition theorem. Consider a particular input quantity X^i. Then,

$$V(Y) = V(E(Y|X^i)) + E(V(Y|X^i)) \tag{2.14}$$

The above variance decomposition is true if and only if there exists a value of Y for

every value $\boldsymbol{X}$; in other words, the G is a deterministic transfer function, as explained at the beginning of this section.

In order to compute the sensitivity of a particular input quantity X^i, this input quantity is first fixed at a particular deterministic value and the expectation of the model output is calculated by considering the variation in other output quantities (denoted by X^{-i}). Thereby, the effect of the uncertainty of all other input quantities is averaged. Then, different deterministic values of the input quantity X^i are considered based on their probability distributions and the variance of the expectation is calculated. This metric is known as the first-order effect index of the input variable X^i on the variance of the output Y:

$$S_1^i = \frac{V_{X^i}(E_{X^{-i}}(Y|X^i))}{V(Y)} \tag{2.15}$$

The first-order effect measures the contribution of the variable X^i by itself. The sum of first order indices of all variables is always less than or equal to unity. The difference between this sum and unity is representative of the interaction between the input variables. Further, higher the first-order effect, more important the variable is.

The interaction or combined effect of two variables X^i and X^j can also be calculated similarly. Now, both these input quantities are fixed at particular deterministic values (denoted by $X^{i,j}$) and the expectation of the model output is calculated by considering the variation in other output quantities (denoted by $X^{-i,j}$). Thereby, the effect of the uncertainty of all other input quantities is averaged. Then, different deterministic values of both X^i and X^j are considered based on their probability distributions and the variance of the expectation is calculated as:

$$S_1^i + S_1^j + S_2^{i,j} = \frac{V_{X^{i,j}}(E_{X^{-i,j}}(Y|X^{i,j}))}{V(Y)} \tag{2.16}$$

The expression in Eq. 2.16 accounts not only for the individual effects of X^i and X^j but also for the interaction between X^i and X^j. The term $S_2^{i,j}$ is called as the second-order index, which explains only the interaction between X^i and X^j. Similarly the third-order effects, fourth-order effects, etc. can also be calculated. These quantities are collectively known as variance-based sensitivity indices or Sobol's indices [36, 37].

If there are n input quantities, then there are n first-order effects, ${}^nC_2$ second-order effects, ${}^nC_3$ third-order effects, and so on until ${}^nC_n$ n^{th} order effects. Hence this requires the calculation of $2^n - 1$ indices, which may be computationally intensive. Hence, researchers often compute only the first-order index and the so-called total effects index for each input quantity.

Consider the expression $\frac{V_{X^{-i}}(E_{X^i}(Y|X^{-i}))}{V(Y)}$. In analogy with the above discussion regarding Eq. 2.16, this expression includes all interaction terms of all orders concerning all variables X^{-i}; any term involving X^i (both individual and any interaction with others) would not be included. As the sum of all the sensitivity indices must be equal to unity, the total effects (the sum of individual effects of X^i and all interactions with other quantities) can be calculated as:

$$S_T^i = 1 - \frac{V_{X^{-i}}(E_{X^i}(Y|X^{-i}))}{V(Y)} \tag{2.17}$$

The sum of the total effects indices of all variables is always greater than or equal to unity; equality holds when there is no interaction between the input quantities. (In this case, the first-order effects indices are equal to the total effects indices). If the total effects index is low, then it means that the input quantity is not important.

(In the rest of the dissertation, for the sake of simplicity, the domain for expectation and variance are not explicitly written. In other words, $V_{X^i}(E_{X^{-i}}(Y|X^i))$ and $V_{X^{-i}}(E_{X^i}(Y|X^{-i}))$ are replaced with $V(E(Y|X^i))$ and $V(E(Y|X^{-i}))$ respectively).

It is important to calculate both the first-order effects and the total effects indices.

If the first-order index of a particular variable X^i is low, then it is not necessary that this variable is unimportant. The interaction of this variable with other variables may contribute significantly to the variance of Y and hence, there is a possibility that X^i is, in fact, an important variable. The effects of interaction are reflected in the total effects index. Further, the difference between the total effects index and the first-order effects index provides an estimate of the contribution of variance due to the interaction between X^i and other variables. Thus, both the first-order and total effects indices must be computed in order to assess the sensitivities of the variables. The above discussion only considered uncertain model inputs; if there are additional model parameters (whose uncertainty is known), then the inputs can be appended with the parameters as $Y = G(\boldsymbol{X}; \boldsymbol{\theta})$ and the above analysis can be performed.

The method of global sensitivity analysis will be used in several chapters in this dissertation. In Chapter III, it is used to quantify the individual contributions of physical variability and epistemic uncertainty. In Chapters V and VIII, it is used for screening of calibration parameters. Finally, in Chapter IX, it is used in the development of the new inverse sensitivity analysis methodology, where the focus is to analyze the sensitivity of inverse analysis, i.e. sensitivity of calibration result to the various sources of uncertainty.

2.7 Markov Chain Monte Carlo Sampling

The class of Markov Chain Monte Carlo methods can be used to generate samples from an arbitrary probability distribution, especially when the CDF is not invertible or when the PDF is known only up to a proportionality constant. In Section 2.4, it was explained that the latter is the case in Bayesian inference, where the objective is to compute the posterior distribution. Therefore, MCMC sampling can be used to draw samples from the posterior distribution, and these samples can be used in

conjunction with the kernel density estimation procedure to construct the posterior distribution.

There are several algorithms which belong to the class of MCMC sampling methods. Two such algorithms, the Metropolis algorithm [38] and the slice sampling [39] algorithm are discussed below.

2.7.1 The Metropolis Algorithm

Assume that a function that is proportional to the PDF is readily available, as $f(x)$; this means that $f(x)$ is not a valid PDF because $\int f(x)dx \neq 1$. For the purpose of illustration, consider the one dimensional case, i.e. $x \in \mathbb{R}$. The following steps constitute the algorithm in order to generate samples from the underling PDF. Note that, the function $f(x)$ is always evaluated at two points and the ratio is only considered; the effect of the unknown proportionality constant is therefore nullified.

1. Set $i = 0$ and select a starting value x_0 such that $f(x_0) \neq 0$.
2. Initialize the list of samples $X = \{x_0\}$.
3. Repeat the following steps; each repetition yields a sample from the underlying PDF.
   (a) Select a prospective candidate from the proposal density $q(x^*|x_i)$. The probability of accepting this sample is equal to $\frac{f(x^*)}{f(x_i)}$.
   (b) Calculate acceptance ratio $\alpha = \min\,(1, \frac{f(x^*)}{f(x_i)})$
   (c) Select a random number u, uniformly distributed on [0, 1].
   (d) If $u < \alpha$, then set $x_{i+1} = x^*$, otherwise set $x_{i+1} = x_i$.
   (e) Augment the list of samples in X by x_{i+1}.
   (f) Increment i, i.e. $i = i + 1$.

4. After the Markov chain converges, the samples in X can be used to construct the PDF of X using kernel density estimation.

The common practice is to generate a few hundreds of thousands of samples and discard the first few thousand samples to ensure the convergence of the Markov Chain.

The Metropolis algorithm [38] assumes that the proposal density is symmetric, i.e. $q(x^*|x_i) = q(x_i|x^*)$. A generalization of this algorithm assumes asymmetric proposal density functions $q_1(x^*|x_i)$ and $q_2(x_i|x^*)$; this algorithm is referred to as Metropolis-Hastings algorithm[40]. The only difference is that the probability of accepting the prospective candidate is calculated as $\frac{f(x^*)q_2(x_i|x^*)}{f(x_i)q_1(x^*|x_i)}$.

2.7.2 Slice Sampling

Consider the same function $f(x)$, i.e. the PDF of X, known up to a proportionality constant. The steps of the slice sampling algorithm are as follows:

1. Set $i = 0$ and select a starting value x_0 such that $f(x_0) \neq 0$.

2. Draw a random number y from the uniform distribution $[0,\ f(x)]$.

3. Consider the set $f^{-1}[y, \infty)$; note that this set may not be convex, especially when the target distribution is multi-modal. Select a sample which is uniformly distributed on this set. Assign $i = i + 1$, and call this sample x_i.

4. Repeat Steps 1 - 3 to generate multiple samples of X and construct the PDF of X using kernel density estimation.

In contrast with the previously discussed Metropolis algorithm, the slice sampling algorithm is not a acceptance-rejection algorithm.

2.7.3 MCMC Sampling: Summary

In addition to the above algorithms, other MCMC sampling algorithms such as Gibbs sampling [41], multiple-try Metropolis [42], Metropolis-within-Gibbs [43] etc. are also discussed in the literature. One important contribution of this dissertation is to replace MCMC sampling with quadrature-based integration to accurately estimate the posterior PDF. A numerical integration method is developed in Section 4.3.7 for this purpose. Further, a new updating methodology is developed in Chapter IX and MCMC sampling is completely eliminated during Bayesian inference.

2.8 Gaussian Process Surrogate Modeling

All of the previously discussed uncertainty quantification methods require repeated evaluation of mathematical models, which are often physics-based, and may be computationally expensive. One approach to overcome this computational difficulty is to make use of surrogate models to replace the original physics-based model. A few evaluations of the original model are used to train this inexpensive, efficient surrogate model. Different types of surrogate modeling techniques such as polynomial response surface [44], polynomial chaos expansion [45], support vector regression [46], relevance vector regression [47], and Gaussian process interpolation [48–50] have been investigated in the literature. This dissertation uses the Gaussian process (GP) surrogate model whenever it is required to replace an expensive model for efficient uncertainty analysis (in Chapters IV – IX). In Chapter III, the Gaussian process model is also used simply as an interpolation tool to interpolate data.

The Gaussian process interpolation is a powerful technique based on spatial statistics and is increasingly being used to build surrogates to expensive computer simulations for the purposes of optimization and uncertainty quantification [48–50]. The

GP model is preferred in this research because (1) it is not constrained by functional forms; (2) it is capable of representing highly nonlinear relationships in multiple dimensions; and (3) can estimate the prediction uncertainty which depends on the number and location of training data points.

The basic idea of the GP model is that the response values Y evaluated at different values of the input variables $\boldsymbol{X}$, are modeled as a Gaussian random field, with a mean and covariance function. Suppose that there are m training points, x_1, x_2, x_3 ... x_m of a d-dimensional input variable vector, yielding the output values $Y(x_1)$, $Y(x_2)$, $Y(x_3)$ ... $Y(x_m)$. The training points can be compactly written as x_T vs. y_T where the former is a $m \times d$ matrix and the latter is a $m \times 1$ vector. Suppose that it is desired to predict the response (output values y_P) corresponding to the input x_P, where x_P is $p \times d$ matrix; in other words, it is desired to predict the output at n input combinations simultaneously. Then, the joint density of the output values y_P can be calculated as:

$$p(y_P|x_P, x_T, y_T; \Theta) \sim N(m, S) \tag{2.18}$$

where Θ refers to the hyperparameters of the Gaussian process, which need to be estimated based on the training data. The prediction mean and covariance matrix (m and S respectively) can be calculated as:

$$\begin{aligned} m &= K_{PT}(K_{TT} + \sigma_n^2 I)^{-1} y_T \\ S &= K_{PP} - K_{PT}(K_{TT} + \sigma_n^2 I)^{-1} K_{TP} \end{aligned} \tag{2.19}$$

In Eq. 2.19, K_{TT} is the covariance function matrix (size $m \times m$) amongst the input training points (x_T), and K_{PT} is the covariance function matrix (size $p \times m$) between the input prediction point (x_P) and the input training points (x_T). These covariance matrices are composed of squared exponential terms, where each element of the matrix

is computed as:

$$K_{ij} = K(x_i, x_j; \Theta) = -\frac{\theta}{2}[\sum_{q=1}^{d} \frac{(x_{i,q} - x_{j,q})^2}{l_q}] \tag{2.20}$$

Note that all of the above computations require the estimate of the hyper parameters Θ; the multiplicative term (θ), the length scale in all dimensions (l_q, $q = 1$ to d), and the noise standard deviation (σ_n) constitute these hyperparameters ($\Theta = \{\theta, l_1, l_2 \ ... \ l_d, \sigma_n\}$). As stated earlier, these hyperparameters are estimated based on the training data by maximizing the following log-likelihood function:

$$log\ p(y_T|x_T; \Theta) = -\frac{y_T^T}{2}(K_{TT} + \sigma_n^2 I)^{-1} y_T - \frac{1}{2} log|(K_{TT} + \sigma_n^2 I)| + \frac{d}{2} log(2\pi) \tag{2.21}$$

Once the hyperparameters are estimated, then the Gaussian process model can be used for predictions using Eq. 2.19. Note that the "hyperparameters" of the Gaussian process are different from the "parameters" of a generic parametric model (for e.g. linear regression model). This is because, in a generic parametric model, it is possible to make predictions using only the parameters. On the contrary, in the case of the Gaussian process model, all the training points and the hyperparameters are both necessary to make predictions, even though the hyperparameters may have estimated previously. For details of this method, refer to [48–55].

An important issue in the construction of the Gaussian process model is the selection of training points. In general, the training points may arise out of field experiments or may be generated using a computer code. This dissertation only considers the latter case and hence, there is no noise in the data, thereby eliminating σ_n from the above equations. Adaptive techniques can be used to select training points for the GP model, in order to construct the response surface to a desired level of accuracy or precision. Since the GP model is capable of estimating the variance in model output, a variance minimization algorithm proposed by McFarland [52] identifies the next

training point at the input variable value which corresponds to the largest variance. This selection algorithm is repeated and training points are adaptively identified until the estimated variance is below a desired threshold. Alternatively, another training point selection algorithm has been developed by Hombal and Mahadevan [56], where the focus is to select successive training points so that the bias error in the surrogate model is minimized.

Once the training points are selected and the surrogate model is constructed, it can be used for (1) Monte Carlo simulation; (2) Markov Chain Monte Carlo simulation; and (3) global sensitivity analysis. It must be noted that the replacement of a complex computer simulation with an inexpensive surrogate leads to approximations; therefore, it is important to include the effect of this approximation in the procedure for overall uncertainty quantification. This feature will be discussed in detail in Chapter IV.

2.9 Summary

This chapter discussed certain essential topics which are fundamental to the development of the new methods in this dissertation. First, the fundamentals of probability theory and the two interpretations of probability were introduced. The philosophy of Bayesian subjective probabilities will be used in the majority of this dissertation, in order to facilitate integration of various sources of uncertainty across multiple levels of models and data. The Bayesian network will be used for such integration, and employed for both uncertainty propagation (forward problem) and inference (inverse problem). Further, this chapter also discussed methods for forward uncertainty propagation (including global sensitivity analysis) and inference (Markov Chain Monte Carlo sampling). While the former will be used repeatedly in the remainder of this

dissertation, new alternative methods for inverse analysis are explored in the forthcoming chapters of this dissertation.

CHAPTER III

DATA UNCERTAINTY: SPARSE AND INTERVAL DATA

3.1 Introduction

In engineering design problems, it is often required to calculate the system response function $Y = G(\boldsymbol{X})$ as a function of the input variables $\boldsymbol{X}$. Several of these input variables $\boldsymbol{X}$ might be uncertain in nature, and this in turn, leads to uncertainty in the output Y. It is important to quantify the uncertainty in the system response function, based on the input uncertainty. This is commonly referred to as uncertainty propagation; since the focus is on quantifying the uncertainty in the output of a model, the term "forward uncertainty quantification" is used in this dissertation.

In the problem of forward uncertainty quantification, probability distributions are assigned to all the uncertain input variables and the probability distribution of the system response function is computed using uncertainty propagation methods such as Monte Carlo sampling [28], first order and second order reliability methods (FORM and SORM respectively) [27, 32]. A brief overview of these methods was provided in Section 2.5.

Though these uncertainty propagation methods have received much attention in the reliability literature, there are two major challenges in the application of such methods to practical engineering applications. The first challenge lies in the characterization of input uncertainty. The second challenge lies in the choice of the model (G) which is used to compute the response; no model is an accurate representation of reality, and hence it is also necessary to quantify the uncertainty in the the model.

The current chapter focuses on the first challenge, i.e. characterizing input uncertainty based on available data, especially when the data is in the form of sparse point and interval data. Chapter IV addresses the issue of model uncertainty in detail.

3.2 Challenges and Existing Approaches

Consider a particular input quantity X. In order to characterize (represent) the input uncertainty, data on that particular quantity needs to be available. Conventionally, probability distributions have been constructed using only point data (e.g., [27]).

Sometimes sufficient point data may not be available to construct such probability distributions. This problem is further complicated if there are interval data. There are several sources of interval data in engineering applications [57–59]. Sometimes, the only information available might come from physical and theoretical constraints that impose bounds on the quantities of interest. Data collected based on temporally spaced inspections may lead to intervals. Uncertainty and errors associated with calibrated instruments may result in experimental observations that are best described using intervals. Sometimes, subject matter experts may describe uncertain quantities using a range of values. Interval data needs to be treated carefully, especially when there are multiple intervals from different sources (say, from multiple experts) and the width of each interval is comparable to the magnitude of the quantity.

The presence of interval data complicates uncertainty representation and propagation because non-probabilistic interval analysis methods have commonly been used to quantify the uncertainty due to interval data. Sometimes, intervals have been approximated with uniform distributions, based on the principle of maximum entropy [60]. This approach may be suitable if a quantity is represented with only a single interval; if multiple intervals are available for a particular quantity, then this method is not suitable since there may be multiple possible uniform distributions.

Sandia National Laboratories conducted an epistemic uncertainty workshop [57] that invited various views on the quantification and propagation of epistemic uncertainty [58], mostly in the form of interval data. Several researchers published different approaches to tackle interval uncertainty in a special issue of the journal Reliability Engineering and Systems Safety [57]. The following paragraphs briefly discuss the different approaches for the treatment of epistemic uncertainty and develop the motivation for the likelihood-based approach proposed in this chapter.

Most probabilistic techniques rely on the existence of sufficient point values for the stochastic quantity of interest. An empirical distribution function (EDF) can be constructed and popular inference techniques such as least squares, moment matching and maximum likelihood can be used to fit parametric probability distributions. The concept of empirical distribution can be extended to interval data sets to arrive at the so-called empirical p-box [59], which is the collection of all possible EDFs for the given set of intervals. Zaman et al. [61] have used the Johnson family of distributions to represent interval uncertainty using a family of distributions which are bounded by the aforementioned p-box. Similar to frequentist p-boxes, Bayesian p-boxes have also been used to represent uncertainty [62].

Researchers have also investigated the use of non-probabilistic approaches for the treatment of epistemic uncertainty due to interval data. Evidence theory [63, 64] has been proposed to handle interval data. This theory is based on the assumption that the sources of interval data are independent. However, data obtained from different sources needs to be properly aggregated. Dempster's rule [65, 66] is a popular scheme of aggregation used for this purpose; several improved rules have also been proposed that acknowledge the conflicts that can potentially exist among evidences from different sources. Convex models of uncertainty [67, 68] use a family of convex functions to represent realizations of uncertain quantities and this approach has also

been used to handle interval data. Zadeh's Extension Principle [69] can be used to construct the possibility distribution of an interval variable which can then be used for uncertainty representation and propagation. Rao and Annamdas [70] presented the idea of weighted fuzzy theory for intervals, where fuzzy set-based representations of interval variables from evidences of different credibilities are combined to estimate the system margin of failure.

The above mentioned methods for uncertainty quantification can be computationally expensive in the context of uncertainty propagation. The application of probabilistic techniques to interval data is computationally expensive and too cumbersome to apply without severe restrictions because there is an infinite number of possible empirical distributions bounded by the p-box. On the hand, non-probabilistic techniques are interval analysis-based approaches [64] and are computationally expensive wherein the cost increases exponentially with the number of uncertain variables, and with the increase in non-linearity of the response function that depends on these uncertain variables. Suppose that some variables are described using intervals and some other physically variable (aleatory) quantities are described using probability distributions. For every combination of interval values, the probabilistic analysis for the aleatory variables has to be repeated, resulting in a computationally expensive double-loop (second-order) sampling analysis.

The primary motivation of this chapter is to accurately perform uncertainty propagation in the presence of sparse point and/or interval data and simultaneously reduce the computational effort involved in uncertainty propagation. This is facilitated by the inherent use of the Bayesian philosophy, which allows probabilistic representation of epistemic variables which are described using intervals. (Recall that a frequentist approach will not permit the assignment of probability distributions to epistemic variables.) Therefore, the computational expense can be reduced by including both

aleatory and epistemic variables in a single analysis loop, and well-known probabilistic methods are alone sufficient for uncertainty propagation.

However, it is not straightforward to construct probability distributions in the presence of sparse point and/or interval data. If there are sufficient point values, then both parametric (where the parameters are estimated using the method of moments [27]) and non-parametric (using kernel density estimation [29]) probability distributions can be constructed based on the point values. In the presence of sparse point data, there is large uncertainty in the parameter estimates which need to be accounted for; this issue is explained in detail in the next subsection. In the presence of interval data, it is not easy to calculate moments or construct kernels; hence the construction of a probability density function (PDF) is not straightforward. In order to overcome these challenges, this chapter proposes a new likelihood-based methodology to facilitate probabilistic representation of quantities described using sparse point and/or interval data.

3.3 Proposed Approach

Consider an input quantity which needs to be represented using a probability distribution. Conventionally, a distribution type (e.g., normal, lognormal, etc.) is *assumed* for the quantity of interest, and the parameters of this probability distribution (e.g., mean and standard deviation in the case of a normal distribution) are usually estimated using techniques of statistical inference using observed data. When large amount of point data is available, it may be reasonable to compute deterministic estimates of the parameters using techniques such as the the method of maximum likelihood [22], method of moments [27], etc. It can be proved that, under some conditions, these deterministic estimates approach the true estimates as the number of data approaches infinity [71]. When sparse and/or data is available, there is uncertainty

associated with these estimates; the importance of this uncertainty in the distribution parameters increases especially when the data is sparse and/or imprecise. The topic of distribution parameter uncertainty has been studied by several researchers in the past [72–74] and this has also been referred to as statistical uncertainty [27, 75] or second-order uncertainty [76, 77]. Further, the distribution type (which needs to initially assumed) may be uncertain, and therefore, it is also necessary to quantify the uncertainty due to the distribution type.

Hence, in general, the uncertainty in the quantity of interest is composed of three parts:

1. Choice of distribution type
2. Choice of distribution parameters
3. Given the choice of distribution type and distribution parameters, the uncertainty in the variable is due to the physical variability; for the sake of simplicity, in the rest of chapter, this is simply referred to as variability.

The proposed likelihood-based methodology is developed for three different cases, as follows:

1. **Case 1**: The distribution type of the variable X is assumed to be known and the distribution parameters are estimated. This is a parametric approach (Section 3.4).
2. **Case 2**: The distribution type of the variable X is unknown; however, it is assumed that multiple competing PDF types are available. Each of the competing PDF types are parametric PDFs and therefore, this is also a parametric approach. The uncertainty in the distribution type and the distribution parameters are both estimated (Section 3.6).

3. **Case 3**: The distribution type of the variable X is unknown and no assumption is made regarding the PDF type. Interpolation techniques are used to estimate the PDF and therefore, this is a non-parametric approach (Section 3.7).

In addition to the above cases, a new computational methodology is developed to quantify the individual contributions of variability and epistemic uncertainty (distribution type uncertainty and distribution parameter uncertainty). This methodology is based on sensitivity analysis and is developed for Cases 1 and 2 above (in Sections 3.5 and 3.6 respectively). The sensitivity analysis methodology cannot be applied to Case 3, because the non-parametric PDF accounts for all the three types of uncertainty - variability, distribution type uncertainty, and distribution parameter uncertainty.

In each of the above sections, simple numerical examples are illustrated. Finally, the proposed methods are illustrated using the challenge problem from Sandia Epistemic Uncertainty Workshop [57] in Section 3.8.

3.4 Case 1: Known PDF Type (Parametric)

3.4.1 Estimation of Distribution Parameters

This section considers the case where the distribution type of a particular quantity X is known. Let the corresponding PDF be denoted by $f_X(x|\boldsymbol{P})$, where $\boldsymbol{P}$ denotes the distribution parameters which need to be estimated based on the available data. First, the concept of likelihood is reviewed for point data and then extended to the case of interval data.

Assume that x_i ($i = 1$ to m) point valued data are available for the estimation of $\boldsymbol{P}$. The likelihood function of the parameters, denoted by $L(\boldsymbol{P})$ is defined as being proportional to the probability of observing the given data (x_i; $i = 1$ to m)

conditioned on the parameters P [19, 22]. Note that, by definition, the definition of likelihood is meaningful only up to a proportionality constant [19].

Given $\boldsymbol{P}$, X follows a continuous PDF $f_X(x|\boldsymbol{P})$. For a continuous density function, the probability value for any single discrete point x_i is theoretically zero. Hence, Pawitan [22] states, "A slight technical issue arises when dealing with continuous outcomes, since theoretically the probability of any point value x_i is zero. We can resolve this problem by admitting that in real life, there is only a finite precision: observing x_i is short for observing $X \in (x_i - \frac{\epsilon}{2}, x_i + \frac{\epsilon}{2})$ where ϵ is the precision limit." If ϵ is small enough, on observing x_i, the likelihood for $\boldsymbol{P}$ is:

$$\begin{aligned} L(\boldsymbol{P}) &\propto P(X \in (x_i - \tfrac{\epsilon}{2}, x_i + \tfrac{\epsilon}{2})|\boldsymbol{P}) \\ &= \textstyle\int_{x_i-\frac{\epsilon}{2}}^{x_i+\frac{\epsilon}{2}} f_X(x|\boldsymbol{P})dx \\ &= \epsilon f_X(x_i|\boldsymbol{P}) \text{ (by mean value theorem)} \\ &\propto f_X(x_i|\boldsymbol{P}) \end{aligned} \tag{3.1}$$

Hence, the likelihood of the parameters $\boldsymbol{P}$ can be calculated as being proportional to the PDF $f_X(x|\boldsymbol{P})$ evaluated at the observed data point. Note that this density function is actually conditioned on the parameters $\boldsymbol{P}$. If there are several data points (x_i; $i = 1$ to m) that are independent of each other, then the combined likelihood of the parameters $\boldsymbol{P}$ can be calculated as:

$$L(\boldsymbol{P}) \propto \prod_{i=1}^{m} f_X(x_i|\boldsymbol{P}) \tag{3.2}$$

(The word "independent" above implies that the sources of data, i.e. different experiments, or different experts from which the data originate, are considered to be statistically independent. In other words, the outcome of one experiment (or the

information from one expert) does not affect the outcome of another experiment (or the information from another expert)).

The parameters $\boldsymbol{P}$ can be inferred by maximizing the expression in Eq. 3.2. This estimate is popularly known as the maximum likelihood estimate.

Note that the above derivation of likelihood considers an infinitesimally small interval around the data point x_i. Hence, it is straightforward to extend this definition to any general interval $[a, b]$. Hence, the expression for likelihood of the parameters P, for a single interval $[a, b]$ is:

$$\begin{aligned} L(\boldsymbol{P}) &\propto P(\text{ data } |\boldsymbol{P}) \\ &= P(X \in [a, b]|\boldsymbol{P}) \\ &= P(a \leq X \leq b|\boldsymbol{P}) \end{aligned} \tag{3.3}$$

The expression in Eq. 3.3 is evaluated from the cumulative distribution function (CDF) $F_X(x|\boldsymbol{P})$, i.e. the integral of the PDF. Therefore,

$$L(\boldsymbol{P}) \propto \int_a^b f_X(x|\boldsymbol{P})dx = F_X(b|\boldsymbol{P}) - F_X(a|\boldsymbol{P}) \tag{3.4}$$

The expression on the right hand side of Eq. 3.4 represents the area under the PDF, and hence its value is always less than or equal to unity.

In principle, the PDF $f_X(x|\boldsymbol{P})$ can assume any arbitrary shape, the only restrictions being that it should be positive and the area under the curve must be equal to unity. Consider the optimization problem of maximizing the expression on the right hand side of Eq. 3.4. We know that the maximum value of this expression will be equal to unity. This expression will be equal to unity for any arbitrary density function $f_X(x)$ whose support is bounded in $[a, b]$ (i.e. $f_X(x) = 0$ when $x < a$ and $x > b$, and $f_X(x)$ takes arbitrary positive values in $a \leq x \leq b$ so that the area the under the

curve is equal to unity) and hence, this optimization problem (i.e. inferring the distribution parameters $\boldsymbol{P}$ for a single interval) will have a non-unique solution. Hence, the likelihood-based approach cannot be used to represent a single interval using a probability distribution.

If there are several intervals (given by $[a_i, b_i]$, $i = 1$ to n) for the description of X and these intervals are assumed to be independent (similar to calculation of likelihood for point data), then the combined likelihood of the parameters $\boldsymbol{P}$ of the PDF $f_X(x|\boldsymbol{P})$ can be calculated similar to Eq. 3.2 as:

$$L(\boldsymbol{P}) \propto \prod_{i=1}^{n} \int_{a_i}^{b_i} f_X(x|\boldsymbol{P}) dx \tag{3.5}$$

If the available data is a combination of both point values and intervals, the likelihood function of the parameters $\boldsymbol{P}$ can be calculated as follows. Suppose there is a combination of point data (m data points, x_i, $i = 1$ to m) and interval data (n intervals, $[a_i, b_i]$, $i = 1$ to n) for the description of X. The right hand sides of Eqs. 3.2 and 3.5 are quantities that are proportional to probabilities that can in turn be multiplied to calculate the combined likelihood. The multiplication is justified by the assumption that the sources of these data are independent, as:

$$L(\boldsymbol{P}) \propto [\prod_{i=1}^{m} f_X(x = x_i|\boldsymbol{P})][\prod_{j=1}^{n} \int_{a_j}^{b_j} f_X(x|\boldsymbol{P}) dx] \tag{3.6}$$

The maximum likelihood estimate of the parameters $\boldsymbol{P}$ can be calculated by maximizing the expression in Eq. 3.6 when both point data and interval data are available. Instead of maximizing likelihood, this dissertation will use a *full likelihood* estimate in which the entire likelihood function is used to construct the PDF of the distribution parameters $\boldsymbol{P}$ as in Eq. 3.7. The idea that the entire likelihood function, rather than

merely its maximizer, should be used for inference was emphasized by Barnard et al. [78].

Let $f(\boldsymbol{P})$ denote the joint PDF of the distribution parameters $\boldsymbol{P}$. Apply Bayes theorem and choose uniform prior density ($f'(\boldsymbol{P}) = h$), and calculate the joint PDF as:

$$f(\boldsymbol{P}) = \frac{hL(\boldsymbol{P})}{\int hL(\boldsymbol{P})d\boldsymbol{P}} = \frac{L(\boldsymbol{P})}{\int L(\boldsymbol{P})d\boldsymbol{P}} \tag{3.7}$$

Note: The uniform prior density function can be defined over the entire admissible range of the parameters $\boldsymbol{P}$. For example, the mean of a normal distribution can vary in $(-\infty, \infty)$ while the standard deviation can vary in $(0, \infty)$ because the standard deviation is always greater than zero. Both these prior distributions are improper prior distributions because they do not have finite bounds.

The construction of the likelihood function in the case of interval data has been considered in the literature (for e.g. [79]). The contribution of this chapter is to use this idea to construct PDFs for such interval data, and hence use such PDFs for uncertainty propagation.

3.4.2 Family of PDFs

The purpose of representing the variable X using probability distributions is to make use of these distributions in the context of uncertainty propagation and reliability evaluation. In this analysis, the variable X then serves as an input to a computational model (that calculates the system output Y), and the aim is to calculate the PDF of the model output Y. The model is evaluated for a sample of X and the corresponding sample model output is calculated. If there is only one PDF for each input X, then this sampling technique can be repeated for different samples of inputs in a Monte Carlo-based technique, and the PDF of Y can be calculated.

However, in the current problem, the variable X is represented by a family of

distributions. Each PDF of X can be used to estimate a corresponding PDF for Y, thereby resulting in a family of estimated distributions for the model output Y. In principle, this requires a double loop Monte Carlo analysis, often called as second-order Monte Carlo analysis [76, 77] . In the outer loop, a sample of $\boldsymbol{P}$ is drawn; this sample determines one distribution of the above-mentioned family. Thus, there is variability in X for every realization of $\boldsymbol{P}$. In the inner loop, several samples of X are drawn from this distribution which is used to calculate one PDF for the model output Y. If each level of sampling generated 1000 members, then this approach would require a total of 10^6 model evaluations to estimate the family of output distributions. This approach is computationally expensive and may not be affordable in may cases.

3.4.3 Unconditional PDF : Predictive Posterior

Alternative to the family of PDFs approach, a single unconditional PDF of X, which includes both the variability in X and the uncertainty in the distribution parameters $\boldsymbol{P}$, as:

$$f_X(x) = \int f_X(x|\boldsymbol{P}) f(\boldsymbol{P}) d\boldsymbol{P} \tag{3.8}$$

Note that the RHS of Eq. 3.8 is not conditioned on $\boldsymbol{P}$ anymore. Some researchers refer to this PDF $f_X(x)$ as the predictive PDF [72, 75] of X.

There are three possible ways to interpret Eq. 3.8, all of which are valid. The first interpretation is based on conditional probability and total probability. Note that the PDF $f_X(x|\boldsymbol{P})$ is actually a conditional density function, conditioned on the parameters $\boldsymbol{P}$. The unconditional probability density function $f_X(x)$ can be calculated by integrating over the space of the parameters $\boldsymbol{P}$, as shown in Eq. 3.8.

The second interpretation is based on computing the expected value of an arbitrary function $h(\beta)$ which is defined as a function of a random variable β with PDF $f(\beta)$. The expression for the expected value of the function is well-known and can be written

as:

$$E(g(\beta)) = \int h(\beta) f(\beta) d\beta \tag{3.9}$$

Note that $E(g(\beta))$ is independent of β. In the context of distribution parameter uncertainty, $\boldsymbol{P}$ corresponds to the parameters β, and the PDF value at any x, i.e. $f_X(x)$ corresponds to the arbitrary function h. Note that while $h(\beta)$ is a function of β, the PDF $f_X(x|\boldsymbol{P})$ value is a function of the parameters $\boldsymbol{P}$. Therefore, Eq. 3.8 estimates the expected value of $f_X(x)$ similar to Eq. 3.9 which estimates the expected value of h, thus replacing a family of PDFs by an "averaged" PDF.

The third interpretation is that Eq. 3.8, fundamentally, is a weighted sum of $f_X(x|\boldsymbol{P})$ over all possible P's where the weights are given by $f(\boldsymbol{P})$. In other words, this is a mixture of distributions. For example, if $f_X(x|\boldsymbol{P})$ were a normal PDF to start with, then Eq. 3.8 corresponds to a Gaussian mixture.

Eq. 3.8 needs to be evaluated numerically and the resultant PDF is not analytical. A single loop sampling procedure (as against the double loop approach discussed earlier in Section 3.4.2) can be used to generate samples directly from the predictive posterior distribution, as follows:

1. Select <u>one</u> random sample of parameters $\boldsymbol{P}$ based on the PDF $f(\boldsymbol{P})$.

2. Generate <u>one</u> random sample of X based on the PDF $f_X(x|\boldsymbol{P})$, where $\boldsymbol{P}$ was selected in Step 1.

3. Repeat steps 1 and 2 to generate multiple samples of X, and use kernel density estimation to construct the predictive PDF $f_X(x)$.

This single loop sampling technique will be used again in Chapter IX.

Though the variable X was initially assumed to follow a particular type of a parametric PDF $f_X(x|\boldsymbol{P})$, say normal, lognormal, etc., the unconditional, predictive PDF $f_X(x)$ (after the integration in Eq. 3.8) is non-parametric, i.e. the resultant

probability distribution is not of the same type and cannot be classified as normal, lognormal, etc. For example, a mixture of Gaussian distributions is not Gaussian [80].

The integration in Eq. 3.8 does not involve the calculation of any performance function or a computational model (which is usually the more expensive calculation in practical problems of uncertainty propagation or reliability analysis). The integration in Eq. 3.8 handles input uncertainty (combining both distribution parameter uncertainty and variability) before the "uncertainty propagation" stage, thereby leading to a single PDF for X and thereby allows the two loops of sampling to be collapsed into a single loop for the sake of faster computation, especially in the context of uncertainty propagation analysis. A comparison of the family of distributions and the unconditional distribution is shown, through PDFs in Fig. 3.1 and through CDFs in Fig. 3.2.

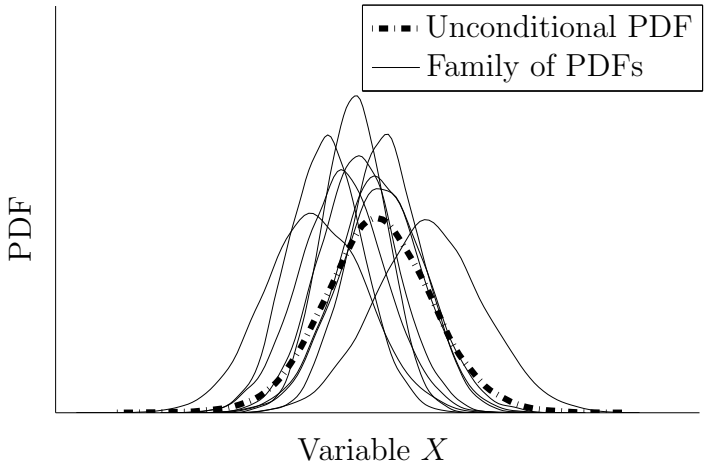


Figure 3.1: Family of Distributions Vs. Single Distribution: PDF

3.4.4 Illustrative Example : Uncertainty Representation

Consider an example where data about X is available in two forms - point data and interval data. Say, 3 point estimates are available as [4.1, 5.6, 3.8] and three sets

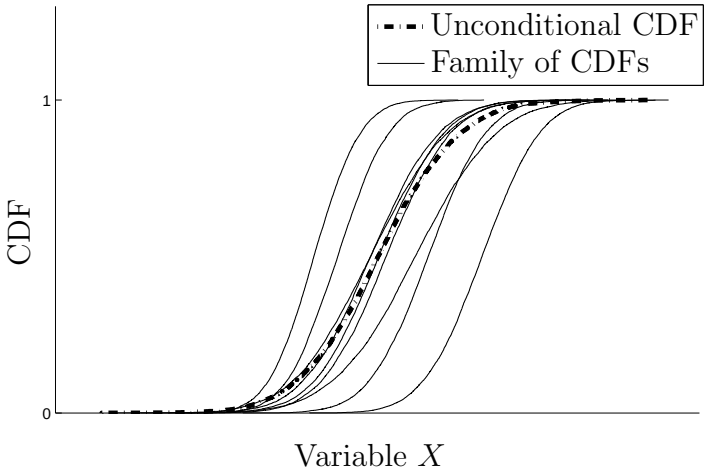
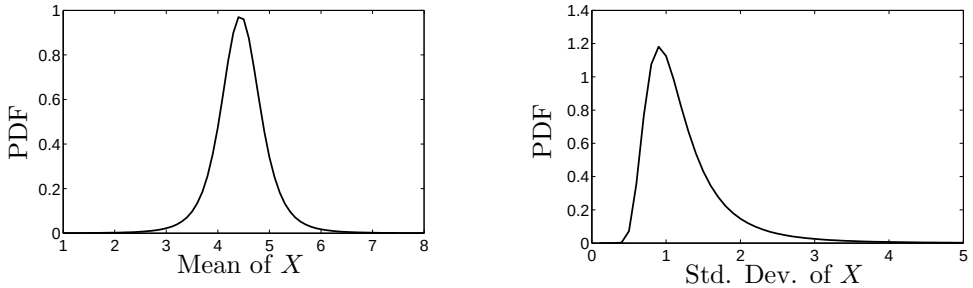


Figure 3.2: Family of Distributions Vs. Single Distribution: CDF

of intervals are available as [3.5, 4], [3.9, 4.1], and [5, 6]. The goal is to represent X using a probability distribution. Once the corresponding unconditional PDF is estimated, it is easy to propagate this uncertainty through a system model.



(a) PDF of Mean (b) PDF of Standard Deviation

Figure 3.3: PDFs of Distribution Parameters

Assume that X is a normal random variable parameterized by mean and standard deviation. Hence, X is normally distributed *conditioned* on the values of mean and the standard deviation. The combined likelihood function considering both point data and interval data is calculated using Eq. 3.6. Note that the likelihood is calculated for the mean and the standard deviation together, i.e. $\boldsymbol{P}$ in $L(\boldsymbol{P})$ denotes the vector of distribution parameters, i.e. the mean and the standard deviation. Hence,

after the joint distribution of the parameters is calculated, the individual (marginal) distributions of the mean and the standard deviation are calculated and shown in Fig. 3.3.

Once the PDFs of the distribution parameters $\boldsymbol{P}$ are calculated, the unconditional PDF of X is calculated using Eq. 3.8, and is plotted in Fig. 3.4. Note that Eq. 3.8 needs to consider the joint PDF of the mean and the standard deviation, and not the marginal distributions, so as to account for any statistical dependence between the parameters. As mentioned earlier, the unconditional PDF reflects the natural variability of X as well as the uncertainty in the distribution parameters $\boldsymbol{P}$ due to sparse and imprecise data, and is not normal in shape even though $f_X(x|\boldsymbol{P})$ corresponds to a normal density function.

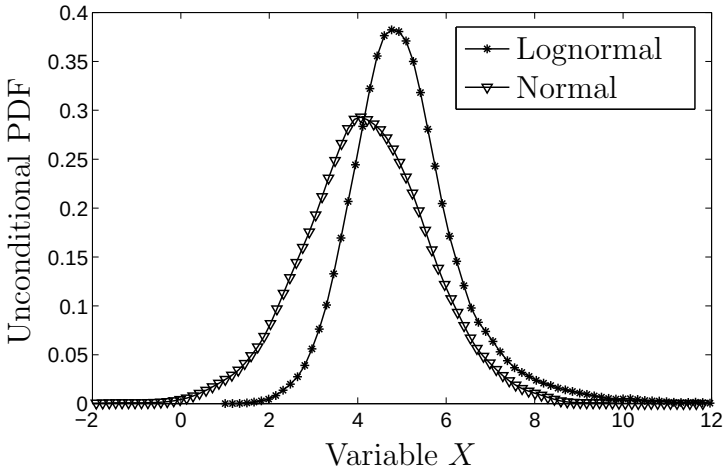


Figure 3.4: Unconditional PDF: Normal vs. Lognormal Assumptions

3.4.5 Remarks

Note that the quantity of interest X was first assumed to follow a particular parametric PDF ($f_X(x|\boldsymbol{P})$) (normal, lognormal, etc.). If the maximum likelihood approach had been used to estimate the most likely value of the parameters $\boldsymbol{P}$, then

parametric form of the distribution (normal, lognormal, etc.) would have been retained. However, after the numerical integration in Eq. 3.8, the resulting unconditional PDF does not have the initially assumed parametric form.

This observation raises the issue of the sensitivity of the final, unconditional PDF to the initial, parametric distribution type assumption. Would different initial parametric distribution assumptions lead to the same final nonparametric distribution? In order to answer this question, the example problem in Section 3.4.4 is repeated assuming that the random variable X follows a lognormal distribution with distribution parameters λ and ξ. The PDFs of λ and ξ are calculated and then these distributions are used to recalculate the unconditional PDF of the random variable X. This result is also shown in Fig. 3.4, along with the result from normal distribution type assumption. It is seen that the distributions resulting from the normal assumption and the lognormal assumption are indeed different. Hence, it is clear that the initial assumption affects the final result of the analysis. This leads to the obvious questions - "What to do if the distribution type is unknown?" and "How to account for distribution type uncertainty?"

There is another important issue to address prior to the treatment of distribution type uncertainty. The use of the predictive posterior distribution combines the effect of variability and distribution parameter uncertainty, in order to aid in fast uncertainty propagation. However, in some applications, it may be necessary to retain the difference between (1) uncertainty in X due to the variability; and (2) uncertainty in X due to uncertainty in the distribution parameters [81]. So, Section 3.5 proposes a computational methodology to apportion the overall uncertainty in X to variability and distribution parameter uncertainty. Distribution type uncertainty is discussed later in Sections 3.6 and 3.7.

3.5 Variability versus Parameter Uncertainty

3.5.1 Need for Assessing Individual Contributions

Though the family of distributions approach is computationally expensive for uncertainty propagation, researchers have sometimes preferred this method because it retains the difference between variability and parameter uncertainty. As a result, it is possible to assess the relative contributions of these two types of uncertainty; while the spread of one particular distribution corresponds to variability, the spread of the family corresponds to parameter uncertainty. However, such an approach only gives a graphical, qualitative measure of the relative contributions of the two types of uncertainty. This section shows that the single (unconditional or predictive) distribution approach does not lose the information, and that in fact it is possible to develop a computational approach to quantify the individual contributions of variability and distribution parameter uncertainty, thus providing more information than the family of distributions approach. Two types of problems are considered:

1. P1: Analysis of contributions of variability and distribution parameter uncertainty within a single variable X

2. P2: Analysis of contributions of variability and distribution parameter uncertainty in multiple input variables $\boldsymbol{X}$ to the output Y of a response function $g(\boldsymbol{X})$, i.e. $Y = g(\boldsymbol{X})$.

The challenge is that the method of global sensitivity analysis discussed earlier in Section 2.6 is not directly applicable to the above problems. Recall that, in global sensitivity analysis, a deterministic transfer function from inputs $\boldsymbol{X}$ to Y is assumed to be readily available, and hence decomposition of variance is valid. This is not the case in the aforementioned problems (P1 and P@).

In problem P1, it is desired to quantify the individual contributions of variability and distribution parameter uncertainty in X to the overall uncertainty in X. GSA cannot be applied directly for this problem because (1) distribution parameter uncertainty is represented by the uncertainty in $\boldsymbol{P}$, whereas there is no variable that separately represents the natural variability in X; and (2) there is no explicit deterministic transfer function from variability and $\boldsymbol{P}$ to X.

In Problem P2, it is desired to quantify the individual contributions of variability and distribution parameter uncertainty in $\boldsymbol{X}$ (a collection of multiple Xs) to the overall uncertainty in $Y = G(\boldsymbol{X})$. However, (1) there is no variable that separately represents the variability in X; and (2) though the distribution parameter uncertainty can be represented by the uncertainty in $\boldsymbol{P}$, this cannot become a direct input to "G" because this will lead to a probabilistic output Y for a fixed value of $\boldsymbol{P}$; GSA needs a deterministic transfer function from $\boldsymbol{P}$ to Y

(Note that $\boldsymbol{X}$ denotes a vector of inputs whereas X denotes a particular input. Similarly, $\boldsymbol{P}$ denotes a vector of distribution parameters whereas P denotes a particular distribution parameter.)

Therefore, the extension of global sensitivity analysis to solve problems P1 and P2 is not trivial. This challenge is overcome by the introduction of an auxiliary variable, as explained below.

3.5.2 The Auxiliary Variable Concept

Two challenges need to be overcome before global sensitivity analysis can be performed: (1) it is necessary to explicitly represent the variability in X; and (2) construct an appropriate deterministic transfer function in order to facilitate the application of GSA. First, variability in X needs to be explicitly represented using a variable. Let U_X denote an auxiliary variable that is introduced for this purpose, i.e.

to explicitly represent the natural variability in X, for a given realization of parameters $\boldsymbol{P}$. In other words, for every U_X, there needs to be a unique X and by varying U_X according to some particular distribution, it must be possible to obtain the entire variability in X given by $f_X(x|\boldsymbol{P})$.

Such an auxiliary variable and its probability distribution can be chosen from the fundamental theory of probability, according to which all probabilities lie between 0 and 1, and hence, there is a unique transformation between any arbitrary (but valid) PDF and the uniform distribution on the interval [0, 1]. Hence, the distribution of U_X is uniform on the interval [0, 1] and the "$U_X \rightarrow X$" transformation is given by:

$$U_X = \int_{-\infty}^{X} f_X(\alpha|\boldsymbol{P})d\alpha \tag{3.10}$$

Note that α is simply a dummy variable for integration.

Eq. 3.10 is simply the definition of the CDF. Given U_X, the corresponding value of X can be calculated by inverting the CDF. Thus, the introduction of the auxiliary variable delineates the overall uncertainty in X into two quantities variability represented by U_X (uniformly distributed on [0, 1]) and distribution parameter uncertainty represented by $\boldsymbol{P}$ (whose distribution was calculated in Eq. 3.7 in Section 3.4.1), as seen in Fig. 3.5. The "spread" of X due to U_X corresponds to the variability whereas the "spread" of X due to $\boldsymbol{P}$ corresponds to distribution parameter uncertainty.

These two quantities U_X and $\boldsymbol{P}$ can be used in the context of Monte Carlo simulation, to generate both the family of distributions for X and the single unconditional (predictive) distribution of X. To generate the family of distributions, first generate one sample of $\boldsymbol{P}$ and then generate several samples of U_X and correspondingly several samples of X (using Eq. 3.10) to generate one member of the family. Then repeat this procedure for multiple samples of $\boldsymbol{P}$. On the other hand, the single predictive

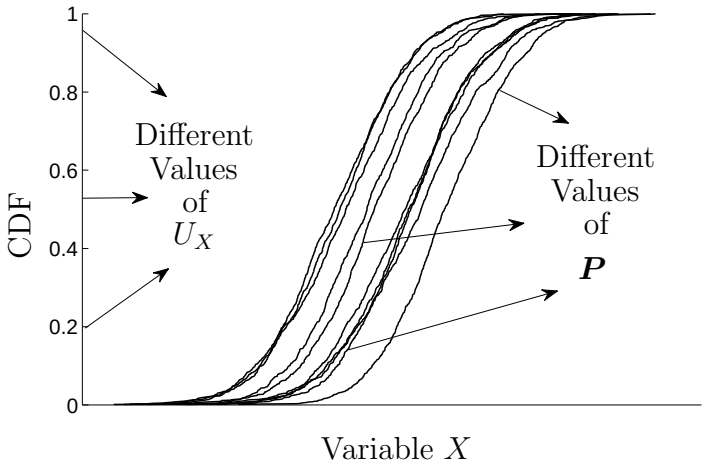


Figure 3.5: Variability (U_X) and Distribution Parameter Uncertainty ($\boldsymbol{P}$)

distribution can be computed by generating a sample of U_X, and an independent sample of $\boldsymbol{P}$ and a corresponding sample of X (using Eq. 3.10), and then repeating the entire procedure in a Monte Carlo sampling framework to generate multiple samples of X and hence a single distribution for X. The fundamental difference is that in the former case, multiple samples of U_X are drawn for each sample of $\boldsymbol{P}$ (double-loop sampling, i.e. sampling U_X within $\boldsymbol{P}$) whereas, in the latter case, one sample of U_X is drawn for each sample of $\boldsymbol{P}$ (single loop sampling, i.e. U_X and $\boldsymbol{P}$ are sampled together, simultaneously and independently).

3.5.3 P1 : Contributions in a Single Variable

To summarize the above development, a variable X with known distribution type but uncertain distribution parameters ($\boldsymbol{P}$) is considered. The variability in X is given by the PDF $f_X(x|\boldsymbol{P})$. The uncertainty in the distribution parameters is represented through the PDF $f(\boldsymbol{P})$, which was calculated in Section 3.4.1. As explained above, now there are two variables that explicitly represent variability (U_X) and parameter uncertainty ($\boldsymbol{P}$); each has its own PDF and the aim is to calculate the contributions of

variability (U_X) and distribution parameter uncertainty ($\boldsymbol{P}$) to the overall uncertainty in X.

Note that there is no computational model here; U_X and $\boldsymbol{P}$ are simply components of X. Therefore, the deterministic transfer function required for GSA is now defined; let "H" denote this function. For the problem at hand, the function "H" is exactly the same as Eq. 3.10, where the "inputs" are U_X and $\boldsymbol{P}$, and the output is X, as shown in Fig. 3.6. In Fig. 6, F^{-1} is inverse of the CDF of X, conditioned on the parameters $\boldsymbol{P}$.

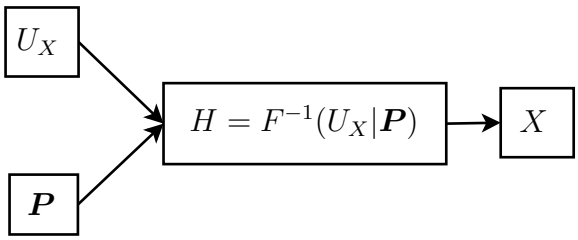


Figure 3.6: Deterministic Transfer Function for GSA

GSA can be now used to calculate the contributions of variability (U_X) and distribution parameter uncertainty ($\boldsymbol{P}$) to the overall uncertainty in X. Hence, using the GSA method in Section 2.6, it is possible to calculate the first-order and total effects of both U_X and $\boldsymbol{P}$. In fact, it is not technically accurate to use the terms "first-order effects" and "total effects" any more. It was explained earlier that these terms are used to assess the effects of a single quantity (X^i in Section 2.6). However, in this case $\boldsymbol{P}$ may be a vector; for example, $\boldsymbol{P}$ consists of two terms, mean and standard deviation, in the case of a normal distribution. Thus in order to calculate the contribution of distribution parameter uncertainty, one must consider the contribution of all variables in $\boldsymbol{P}$ and hence the term "first-order effects" is no more applicable. (In the above example of a normal distribution, calculating the effect of $\boldsymbol{P}$ would in fact require the calculation of second-order effect as in Eq. 2.16.) In order to avoid this

confusion, this dissertation uses the terms "individual effects" and "overall effects" in lieu of "first-order effects" and "total effects". As explained earlier in Section 2.6, it is always necessary to calculate both individual and overall effects.

The individual (I) and overall (O) contributions of variability to the overall uncertainty in X can be calculated as:

$$\begin{aligned} S_U^I &= \frac{V_U(E_{\boldsymbol{P}}(X|U))}{V(X)} \\ S_U^O &= 1 - \frac{V_{\boldsymbol{P}}(E_U(X|\boldsymbol{P}))}{V(X)} \end{aligned} \tag{3.11}$$

The individual (I) and overall (O) contributions of distribution parameter uncertainty to the overall uncertainty in X can be calculated as:

$$\begin{aligned} S_{\boldsymbol{P}}^I &= \frac{V_{\boldsymbol{P}}(E_U(X|\boldsymbol{P}))}{V(X)} \\ S_{\boldsymbol{P}}^O &= 1 - \frac{V_U(E_{\boldsymbol{P}}(X|U))}{V(X)} \end{aligned} \tag{3.12}$$

In the next subsection, this concept is extended where the contributions of variability and distribution parameter uncertainty to a response function are assessed.

3.5.4 P2 : Contributions to a Response Function

Now consider a computational model $Y = G(\boldsymbol{X})$ where $\boldsymbol{X}$ refers to the vector of inputs, and Y refers to the model output. Let X denote a particular input variable. Each input X now has a PDF $f_X(x|\boldsymbol{P})$ (type known), where $\boldsymbol{P}$ refers to the distribution parameters (of X) with PDF $f(\boldsymbol{P})$. Even for a given $\boldsymbol{P}$, note that Y is not deterministic due to the uncertainty in X, and hence GSA cannot be applied directly.

However, in GSA, the aim is to calculate the contribution of each input variable X^i to Y. Here, the aim is different; it is to calculate the contribution of variability and distribution parameter uncertainty in each X to the overall Y. Hence, the uncertainty

in $\boldsymbol{X}$ needs to be decomposed into two parts - variability (U_X) and distribution parameter uncertainty ($\boldsymbol{P}$) - as done in Section 3.5.3.

The deterministic transfer function needed for GSA is constructed with two inputs U_X and $\boldsymbol{P}$ - for each uncertain X. Hence, in $Y = G(\boldsymbol{X})$, each X is replaced with Eq. 3.10 using the corresponding U_X and $\boldsymbol{P}$ as inputs. The output of H is now deterministic; i.e. a single Y for a choice of U_X and $\boldsymbol{P}$ for each X. It is now possible to compute several sensitivity indices regarding the contribution of the following to the variance of Y:

1. Individual and overall effects of the overall uncertainty of any X (by considering the corresponding U_X and $\boldsymbol{P}$ together)

2. Individual and overall effects of variability of any X (by considering the corresponding U_X alone)

3. Individual and overall effects of distribution parameter uncertainty of any X (by considering the corresponding $\boldsymbol{P}$ alone)

4. Individual and overall effects of variability of combinations of multiple Xś, i.e. X^i, X^j, X^k, etc. (by considering the appropriate U^i, U^j, U^k, etc. together)

5. Individual and overall effects of distribution parameter uncertainty of combinations of Xś, i.e. X^i, X^j, X^k, etc. (by considering the appropriate $\boldsymbol{P}^i$, $\boldsymbol{P}^j$, $\boldsymbol{P}^k$, etc. together)

Thus, an easy and efficient methodology is developed for computing the sensitivity indices that show the individual contributions of variability and distribution parameter uncertainty to the overall uncertainty in the model output. Eqs. 3.11 and 3.12 involve the computation of "variance of expectation", which may intuitively require nested loop Monte Carlo sampling; i.e. an inner loop to calculate the expectation and

Table 3.1: Three Cases for Problem P1

Quantity →	Mean (μ)		Standard Deviation (σ)	
Cases	Mean	Standard Deviation	Mean	Standard Deviation
Case 1	10	1	1	0.1
Case 2	10	1	2	0.1
Case 3	10	2	1	0.1

an outer loop to calculate the variance of expectation. It may be argued that this requires the same computational expense as a family of distributions approach. However, there exist single loop sampling approaches to compute these sensitivity indices, as explained in Saltelli et al.[36]. Further, while the family of distributions approach provides only a graphical representation of relative contributions of variability and distribution parameter uncertainty, the proposed approach provides quantitative metrics based on the actual contribution to variance.

The following subsection illustrates the proposed methodology using two examples; first, a single variable with uncertain distribution parameters (problem P1) is considered, and second, the proposed methods are illustrated using an uncertainty propagation (problem P2) problem from the Epistemic Uncertainty Workshop organized by the Sandia National Laboratories [57].

3.5.5 Illustration : Contributions in One Variable

Consider a variable X that is normally distributed with parameters mean (μ) and standard deviation (σ). Both these distribution parameters are assumed to be normally distributed for the sake of illustration. Three different cases are considered, as tabulated in Table 3.1.

For the sake of visualization, the family of PDFs and the single PDF of X are shown for the three cases in Figs. 3.7 - 3.9.

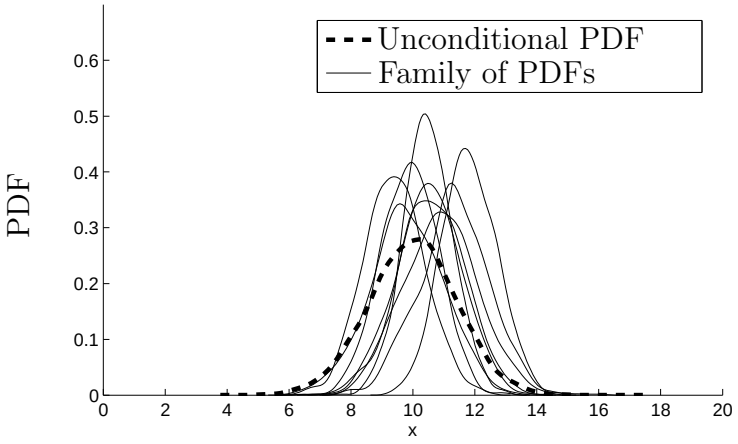


Figure 3.7: Case 1: Family of PDFs and Unconditional PDF

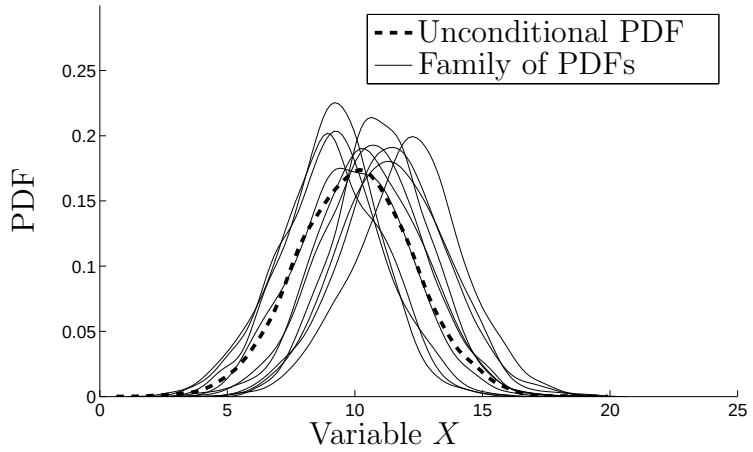


Figure 3.8: Case 2: Family of PDFs and Unconditional PDF

In Eqs. 3.11 and 3.12, $\boldsymbol{P} = [\mu, \sigma]$ is a vector of length two and to calculate the individual effects of $\boldsymbol{P}$, it would actually be necessary to calculate the second-order effects indices, as in Eq. 2.16. The deterministic function is constructed with inputs U_X, μ, and σ. The decomposition of variance is shown in Eq. 3.13.

$$S_U + S_\mu + S_\sigma + S_{U,\mu} + S_{U,\sigma} + S_{\mu,\sigma} + S_{U,\mu,\sigma} = 100\% \tag{3.13}$$

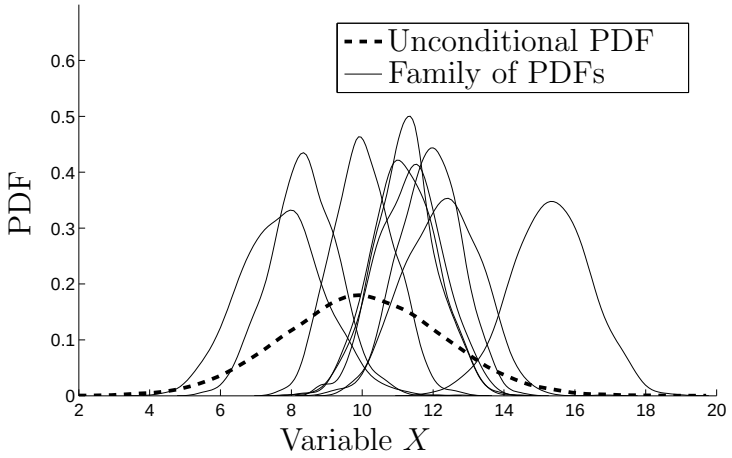


Figure 3.9: Case 3: Family of PDFs and Unconditional PDF

The individual and overall effects of variability and distribution parameter uncertainty are calculated based on Eqs. 3.11 and 3.12. The individual effect of variability is given by S_U; the individual effect of parameter uncertainty is given by $S_\mu + S_\sigma + +S_{\mu,\sigma}$; the overall effect of variability is given by $S_U + S_{U,\mu} + S_{U,\sigma} + S_{U,\mu,\sigma}$; and the overall effect of parameter uncertainty is given by $S_\mu + S_\sigma + S_{U,\mu} + S_{U,\sigma} + S_{\mu,\sigma} + S_{U,\mu,\sigma}$. These sensitivities are tabulated in Table 3.2, in terms of percentage of the total variance.

Table 3.2: Three Cases for Problem P1: Contributions

Cases	**Uncertainty**	**Individual Effects**	**Overall Effects**
Case 1	Variability	50.0 %	50.0%
	Parameter Uncertainty	50.0 %	50.0%
Case 2	Variability	80.0 %	80.0%
	Parameter Uncertainty	20.0 %	20.0%
Case 3	Variability	20.0 %	20.0%
	Parameter Uncertainty	80.0 %	80.0%

The following observations are made from Table 3.2.

1. As seen in Eq. 3.13, the sum of individual effects of variability and the total effects of parameter uncertainty is equal to one. Similarly, the sum of individual

effects of parameter uncertainty and the total effects of variability is equal to one.

2. The contributions of variability and distribution parameter uncertainty are almost equal in Case 1. In Case 2, the mean of the standard deviation is twice as in Case 1 and hence, this increases the contribution of variability and decreases the contribution of distribution parameter uncertainty. In Case 3, the standard deviation of the mean is twice as in Case 1, thereby increasing the distribution parameter uncertainty, and decreasing the contribution of variability, (3) Due to numerical errors that arise due to sampling, all percentage sensitivities are reported only to one decimal place. Though the overall effects indices were greater than the individual effects indices, this is not reflected in Table 3.2 due to the difference being less than 0.1 %. Thus, the overall effects indices are only marginally higher (in fact equal up to 1st decimal place) than the individual effects indices; hence, there is little interaction between the terms corresponding to variability and distribution parameter uncertainty. (4) This analysis suggests that by reducing the contribution of distribution parameter uncertainty, it is possible to reduce the uncertainty in the variable X, for e.g., in Case 1, by approximately 50%. The contribution due to variability is irreducible by collecting more data.

3.5.6 Illustration : Contributions to a Response Function

This section illustrates the proposed sensitivity analysis methodology for quantifying the contributions of variability and distribution parameter uncertainty in input variables (a and b) to the output (y) of a response function $y = (a + b)^a$. Probability distributions (with uncertain distribution parameters) are assumed for inputs a and b. The quantity a is normally distributed with parameters mean (μ) and standard

deviation (σ). The mean is normally distributed as N(0.5, 0.05), and the standard deviation is normally distributed as N(0.1, 0.01). Similarly, the quantity b is normally distributed with parameters mean (α) and standard deviation (β). The mean is normally distributed as N(0.5, 0.05), and the standard deviation is normally distributed as N(0.1, 0.01). The family of distributions for y and the single unconditional distribution are shown in Fig. 3.10.

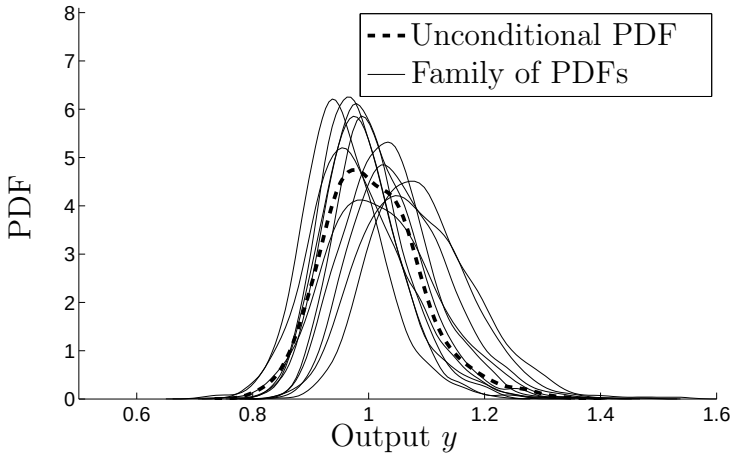


Figure 3.10: Output y: Family of PDFs and Unconditional PDF

Using this information, the objective is to construct the probability distribution of y and apportion the variance in y to variability and distribution parameters of the inputs a and b. Two auxiliary variables, U_A for a and U_B for b are introduced, to represent the variability in each, while the distribution parameter uncertainty is represented by the distributions of μ, σ, α, and β. The deterministic transfer function "H" for GSA is constructed with six inputs: U_A, U_B, μ, σ, α, and β, and the output is y. The variables U_A and U_B are uniformly distributed on [0, 1]. There are 63 terms in the decomposition of variance. Using the GSA method, the following sensitivity indices are calculated. Similar to the previous numerical example, the sensitivity indices are expressed in terms of percentages and both the individual and the overall effects are reported.

Table 3.3: Variability and Distribution Parameter Uncertainty: Contributions

Quantity	**Meaning**	**Individual Effects**	**Overall Effects**
(μ, σ)	Parameter Uncertainty in a	9.0 %	11.8 %
(α, β)	Parameter Uncertainty in b	9.8 %	10.1 %
U_A	Variability in a	37.3 %	43.4 %
U_B	Variability in b	40.3 %	40.3 %
(U_A, U_B)	Total Variability	77.6 %	82.6 %
$(\mu, \sigma, \alpha, \beta)$	Total Parameter Uncertainty	17.4 %	22.4 %

Table 3.3 lists several sensitivity indices the individual and overall effects for (1) distribution parameter uncertainty in a; (2) distribution parameter uncertainty in b; (3) variability in a; (4) variability in b; (5) total variability; and (6) total distribution parameter uncertainty. All sensitivities are reported to one decimal place. Also, the term "quantity" in Table 3.3 refers to the variables which are fixed in the inner loop and outer loop for computation of individual effect and overall effect respectively. Similar to the previous example, it is seen that the interaction between the variability and distribution parameter uncertainty is negligible and that distribution parameter uncertainty contributes to about one-fifth of the overall uncertainty in the output y. Such an analysis clearly helps to identify the important contributors of uncertainty and identify what proportion of the uncertainty can be decreased by collecting more data.

3.5.7 Summary

Section 3.4 developed a method to estimate the distribution parameters in the presence of sparse and interval data; the distribution type was *assumed* to be known. Then an unconditional PDF was computed to combine variability and parameter uncertainty for the sake of uncertainty propagation. Though the PDF combined the two effects, Section 3.5 proposed a computational approach to quantify the individual

contributions of the two types of uncertainty (variability and distribution parameter uncertainty).

In practical applications, the distribution type is seldom known. In Section 3.4.5, it was demonstrated that different assumptions regarding the distribution type leads to different unconditional distributions. This observation motivates the rest of this chapter, which is to focus on the aspect of distribution type uncertainty. Sections 3.6 and 3.7 deal with the distribution type uncertainty. Section 3.6 addresses this issue by considering multiple, competing distribution types while Section 3.7 eliminates the choice of a distribution type by considering non-parametric distributions.

3.6 Case 2: Unknown PDF Type (Parametric)

If the distribution type of a variable X is known, the distribution parameters can be estimated. In many situations, the distribution type itself may not be known and needs to be identified based on the available data. Usually, when adequate data is available, the assumed distribution type can be verified by comparing against empirical distribution functions using statistical goodness-of-fit tests such as chi-square test [82] or Kolmogorov-Smirnov [83] test. Also, Anderson-Darling [84] and Cramer [85] tests are available for multi-variate distributions. In the presence of sparse and interval data, this approach may not be applicable because empirical distribution functions are not unique and are either bounded by a p-box [86] or represented using a family of distributions [61].

In this section, two approaches are pursued for the quantification of model form uncertainty Bayesian model averaging [87, 88], and Bayesian hypothesis testing [89, 90]. The first method is based on assigning weights to competing model forms, and is applicable for comparing two or more models; the weights and the distribution parameters are estimated based on the available data. The Bayesian hypothesis testing

approach computes the extent of support provided by data to the model form and it can be used to assess competing models or to quantify the uncertainty regarding a given model. This approach has been previously used to validate computational models [91] and reliability models [92], and is extended in this dissertation to quantify the uncertainty in the selection of probability distribution type. In addition to distribution type uncertainty, distribution parameter uncertainty is also simultaneously quantified.

Consider a variable X which is an input to a system model. Previously, Sections 3.4 and 3.5 addressed variability and parameter uncertainty in X. When the distribution type is also uncertain, there are three different types of uncertainty: (1) physical variability; (2) distribution type; and (3) distribution parameters. Hence, the total uncertainty about X is equal to the combined effect of all three types of uncertainty. Note that both distribution type uncertainty and distribution parameter uncertainty are examples of epistemic uncertainty, since they can be precisely known if sufficient data were available.

When this input X is propagated through a system model, all three types of uncertainty must be accounted for in a rigorous manner. If a Monte Carlo-based sampling technique is pursued, three loops of sampling are necessary, as shown in Fig. 3.11.

If there is more than one possible distribution type, then the distribution type is treated as a discrete random variable and the distribution type is selected in the first (outermost) loop. In the second (inner) loop, a sample of the distribution parameter value is drawn. These two choices uniquely identify a probability density function for X. In the third (innermost) loop, samples of X are drawn.

Consider for the sake of illustration, a simple case where X could either be Log-normal or Weibull; thus, model form is a discrete variable with two possibilities. If it

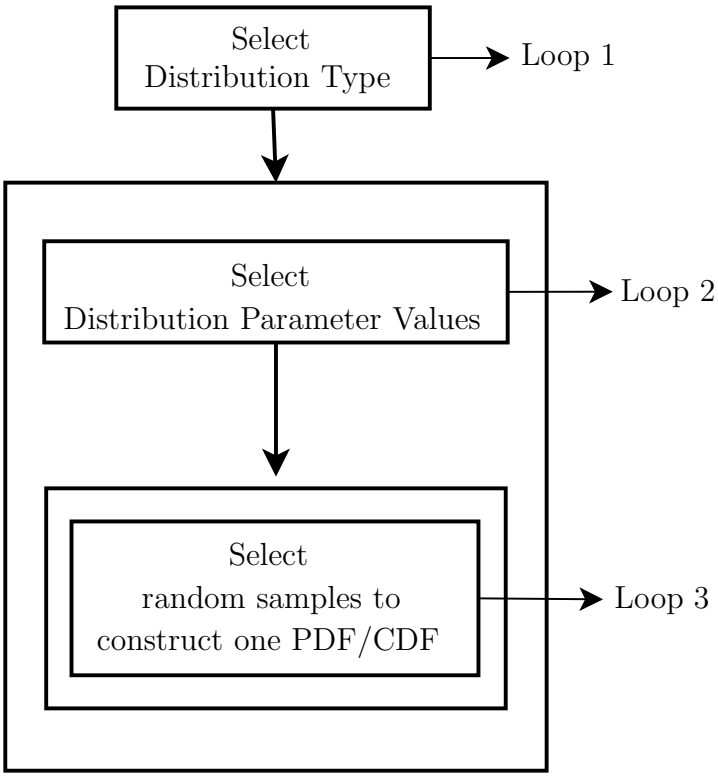


Figure 3.11: Multiple Loops of Sampling

is lognormal, the distribution parameters are mean and standard deviation of $ln(X)$. If it is Weibull, then the parameters are the scale parameter and shape parameter. If these distribution parameters are uncertain, then X can be represented using two families of distributions; the first family consists of Lognormal distributions and the second consists of Weibull distributions. These two families are shown in Fig. 3.12.

The aim is to quantify the extent to which the available sparse/imprecise data support the Lognormal vs. Weibull distribution type assumption. Further, the uncertainty in the distribution parameters of these two distributions also needs to be calculated. First, the method of Bayesian model averaging is investigated in Section 3.6.1, and then the method of Bayesian model averaging is investigated in Section 3.6.2.

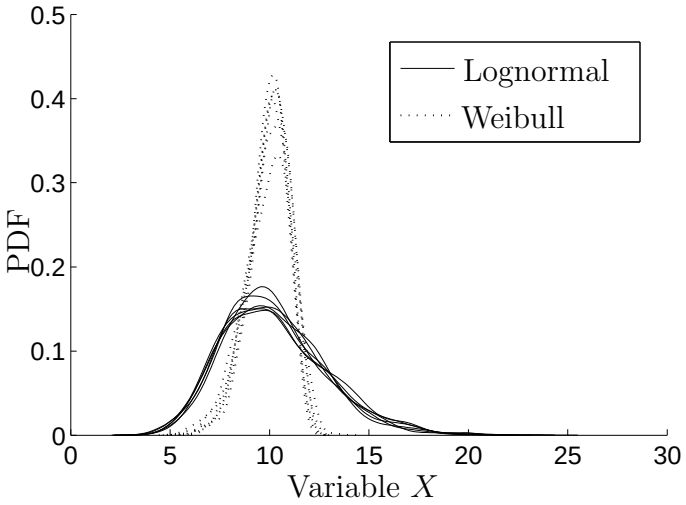


Figure 3.12: Distribution Type and Distribution Parameter Uncertainty

3.6.1 Bayesian Model Averaging Approach

Consider a particular random variable X for which data is available. Let D denote the collection of all data, which comprises of m point data x_i ($i = 1$ to m) and n intervals $[a_i,\ b_i]$ ($i = 1$ to n). Now, the aim is to quantify the model form uncertainty in the probability distribution of X. The method of Bayesian model averaging is applicable when multiple competing model forms are compared. The overall approach is to express the PDF $f_X(x)$ as a weighted sum of the competing model forms.

Without loss of generality, the method is discussed here for two competing model forms; it can be easily extended to any number of competing model forms. Let $f_X^1(x|\boldsymbol{\phi})$ and $f_X^2(x|\boldsymbol{\theta})$ denote the PDFs of the competing model forms; in each PDF, $\boldsymbol{\phi}$ and $\boldsymbol{\theta}$ are the unknown distribution parameters.

Using Bayesian model averaging, the PDF of X can be expressed as the sum of the above two PDFs as:

$$f_X(x|w, \boldsymbol{\phi}, \boldsymbol{\theta}) = w f_X^1(x|\boldsymbol{\phi}) + (1 - w) f_X^2(x|\boldsymbol{\theta}) \tag{3.14}$$

Since there are only two competing distributions here, their weights are chosen as w and $1-w$ in Eq. 3.14. If there are n competing distributions, then there are $n-1$ weights to be estimated, and the n^{th} weight is estimated by imposing the condition that the sum of all the weights is equal to unity, since the area under the PDF $f_X(x|w, \boldsymbol{\phi}, \boldsymbol{\theta})$ must be equal to unity.

A likelihood-based estimation procedure similar to that in Section 3.4.1 is used here. The difference is that the combined likelihood of the weights and the distribution parameters, i.e. $L(\boldsymbol{\phi}, \boldsymbol{\theta}, w)$, is constructed as:

$$L(\boldsymbol{\phi}, \boldsymbol{\theta}, w) \propto [\prod_{i=1}^{m} f_X(x = x_i|w, \boldsymbol{\phi}, \boldsymbol{\theta})][\prod_{j=1}^{n} \int_{a_j}^{b_j} f_X(x|w, \boldsymbol{\phi}, \boldsymbol{\theta})dx] \tag{3.15}$$

This likelihood function can be maximized to obtain the maximum likelihood estimates of $\boldsymbol{\phi}$, $\boldsymbol{\theta}$, and w. Further the uncertainty in the estimates can also be quantified using Bayesian inference, as in Eq. 3.7. A uniform prior bounded on [0, 1] is chosen for w, and non-informative priors are chosen for the distribution parameters $\boldsymbol{\phi}$ and $\boldsymbol{\theta}$. The prior distributions are multiplied by the likelihood function and then normalized to calculate the posterior distributions of $\boldsymbol{\phi}$, $\boldsymbol{\theta}$, and w.

Two illustrations are presented below. The first example considers a large amount of data and two significantly different candidate model forms. The second example considers a large amount of data and two candidate model forms that are <u>not</u> significantly different from one another.

3.6.1.1 Illustration 1

Consider a case of 100 samples generated from an underlying normal distribution with mean and standard deviation equal to 100 units and 10 units respectively. Since the amount of data is large, it is easy to identify that the underlying distribution

is, in fact, normal. However, this example is used only to demonstrate the Bayesian averaging method.

For the sake of illustration, assume that the two competing model forms are normal ($N(\mu, \sigma)$) and uniform ($U(a, b)$). With reference to Eq. 3.15, $\boldsymbol{\phi} = \{\mu, \sigma\}$ and $\boldsymbol{\theta} = \{a, b\}$. Let w denote the weight for the normal distribution, and $1-w$ is the weight for the uniform distribution. The joint likelihood is evaluated for five quantities (w, μ, σ, a, and b), and the posterior distribution is estimated for each quantity using 10,000 samples from slice sampling [39]. The correctness of these posterior distributions can be easily verified since the samples were actually generated from a normal distribution N(100,10).

First, the PDF of the weight w is shown in Fig. 3.13. The estimated statistics/PDFs of distribution parameters are shown in Table 3.4 and Fig. 3.14.

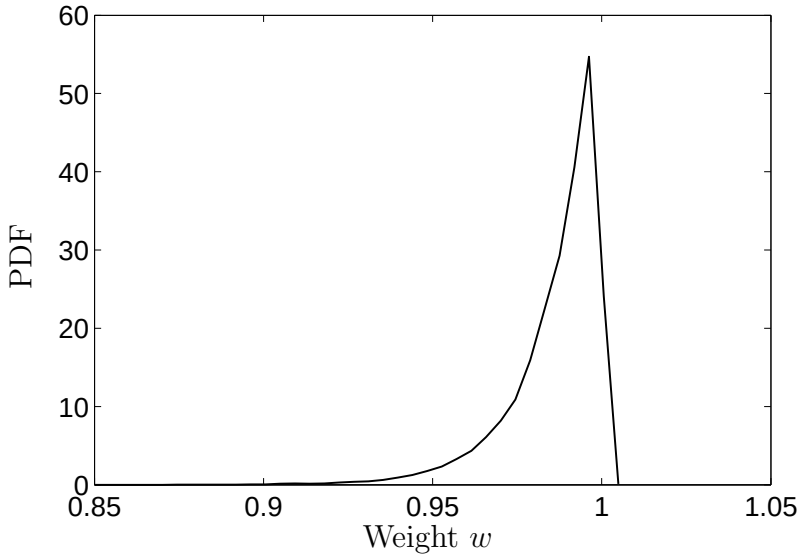
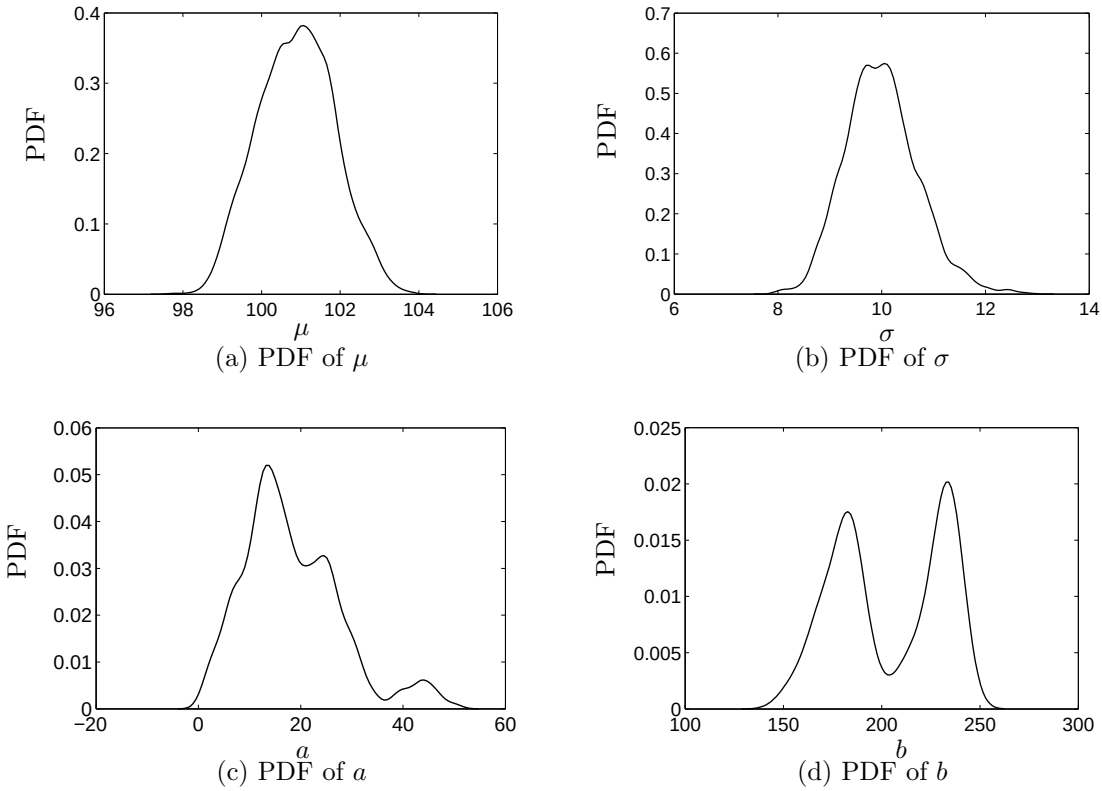


Figure 3.13: PDF of Weight w

From Table 3.4, it can be clearly seen that the method isolates the data to come from a normal distribution. There is high confidence in this conclusion because the mean of w is high (0.98) and the standard deviation of w is small (0.015). Also,



(a) PDF of μ (b) PDF of σ

(c) PDF of a (d) PDF of b

Figure 3.14: PDFs of Distribution Parameters

the distribution parameters of the normal distribution are in good agreement with the actual values using which the data was simulated, and the uncertainty in the estimates of these distribution parameters is small. Since the weighting factor for the uniform distribution is small, the distribution parameter estimates for the uniform distribution have high uncertainty.

The uncertainty in the estimate of the weight w is low because of two reasons: (1) there is sufficient data to conclusively suggest a normal distribution, and (2) the two competing model forms, i.e. normal and uniform are significantly different from each other.

It is obvious that, if there is only sparse data, then the uncertainty in the estimate of w will be high. However, even if there is sufficient data, it is hard to uniquely isolate

Table 3.4: Normal vs. Uniform: Results of Bayesian Model Averaging

Quantity	Mean	Standard Deviation	95% Bounds
w	0.986	0.015	[0.949, 0.999]
μ	100.887	0.969	[99.078, 102.811]
σ	9.998	0.704	[8.752, 11.534]
a	18.193	9.584	[2.997, 43.800]
b	203.767	27.6065	[157.278, 239.324]

a particular model form if the competing model forms are not significantly different from one another, as shown next.

3.6.1.2 Illustration 2

Consider 100 samples generated from an exponential distribution with parameter $\mu = 1$. The PDF for this distribution is given by

$$f_X(x|\mu) = \frac{1}{\mu} exp(-\frac{x}{\mu}) \tag{3.16}$$

For the sake of illustration, assume that the two competing model forms are exponential and Rayleigh. While the former has one parameter (μ) as indicated in Eq. 3.16, the latter also has only one parameter (b), and the PDF is given by:

$$f_X(x|b) = \frac{x}{b^2} exp(-\frac{x^2}{2b^2}) \tag{3.17}$$

Note that the exponential and Rayleigh distributions are not as significantly different from each other as the uniform and normal distributions. This is because both exponential and Rayleigh distributions can be viewed as special cases of the two-parameter Weibull distribution with shape parameters equal to one and two respectively. Since the Weibull distribution is commonly used to study time-dependent reliability, this example is of practical significance.

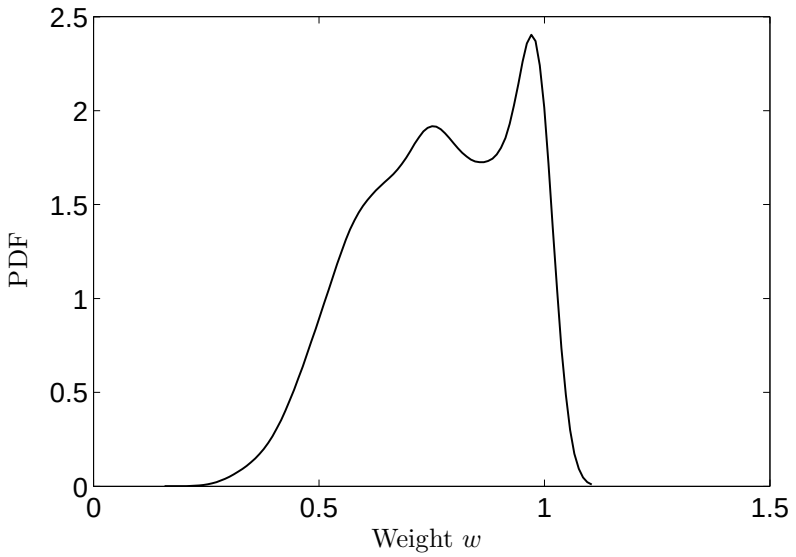


Figure 3.15: PDF of Weight w

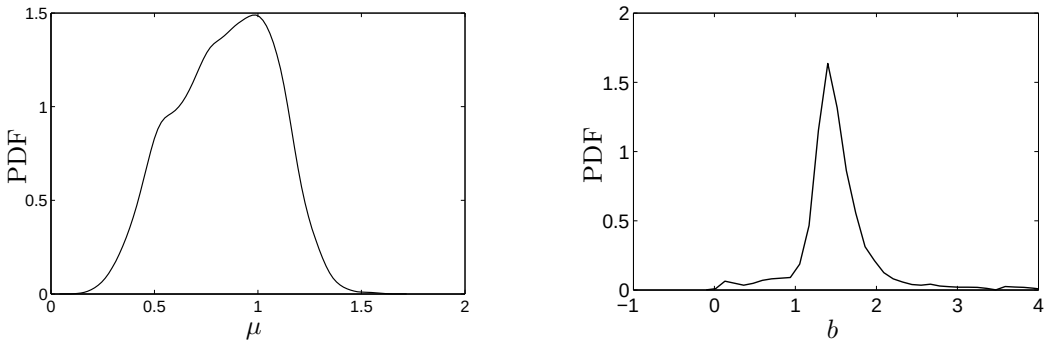
Similar to the previous example, the joint likelihood $L(w, \mu, b)$ is used to evaluate the posterior distributions of w, μ and b respectively. First, the PDF of the weight w is shown in Fig. 3.15 where w is the weight for the exponential distribution, and $1 - w$ is the weight for the Rayleigh distribution.

The PDFs of the distribution parameter for each model-form (μ for exponential distribution and b for Rayleigh distribution) are shown in Fig. 3.16, and the numerical estimates are shown in Table 3.5.

Table 3.5: Exponential vs Rayleigh: Results of Bayesian Model Averaging

Quantity	Mean	Standard Deviation	95% Bounds
w	0.746	0.158	[0.424, 0.988]
μ	0.840	0.239	[0.382, 1.255]
b	2.060	1.181	[0.561,7.793]

The mean of w is about 0.75, which suggests a higher likelihood for the exponential distribution. However, there is significant uncertainty in w, leading to inconclusive distinction between the exponential and Rayleigh distributions. Also, the estimates



(a) PDF of μ (b) PDF of b

Figure 3.16: PDFs of Distribution Parameters

of the distribution parameters suggest a higher likelihood for the exponential distribution, because μ in the exponential distribution has a much smaller uncertainty compared to b in Rayleigh distribution. That is, a "narrow" estimate of μ is sufficient to "explain" the available data whereas a "wide" estimate of b is needed for the same. This is intuitive because the data actually originates from an exponential distribution. Also, the maximum likelihood estimate of μ is one, which is exactly the same as the originally assumed value for μ used to generate the data.

3.6.1.3 Quantifying Individual Contributions

Earlier, Section 3.5 developed a computational method to assess the individual contributions of variability and distribution parameter uncertainty by assuming a particular distribution type. In the present section, the distribution type is also uncertain, and this uncertainty is quantified through the PDF of w. Hence, the method developed in Section 3.5 is now extended to quantify the individual contributions of (1) variability; (2) distribution type uncertainty; and (3) distribution parameter uncertainty.

The concept of the auxiliary variable was introduced earlier in Section 3.5.2 to facilitate the use of global sensitivity analysis for the quantification of individual

Table 3.6: Contributions of Physical Variability and Epistemic Uncertainty

Illustration	Effect	Physical Variability	Distribution Type	Distribution Parameter
Example 1	Individual	94.1%	1.0%	4.0%
Section 3.6.1.1	Overall	98.1%	1.0%	5.2%
Example 2	Individual	40.7%	12.4%	40.5%
Section 3.6.1.2	Overall	43.5%	15.4%	43.3%

contributions. This auxiliary variable is now redefined to include the distribution type uncertainty, as:

$$U_X = \int_{-\infty}^{X} f_X(x|w, \boldsymbol{\phi}, \boldsymbol{\theta})dx \tag{3.18}$$

Similar to Section 3.5, there is unique one-to-one mapping between X and U_X. Since Eq. 3.18 is simply the definition of CDF, by varying U_X on the uniform distribution U(0,1), it is possible to obtain the entire distribution of X. Now, global sensitivity analysis can be applied to calculate the individual and overall effects of physical variability (U_X), distribution type uncertainty (w), and distribution parameter uncertainty ($\boldsymbol{\phi}$ and $\boldsymbol{\theta}$). For example, the results of sensitivity analysis for the illustrative examples discussed in Sections 3.6.1.1 and 3.6.1.2 are tabulated in Table 3.6.

Similar to Section 3.5, it is straightforward to quantify the contributions in a single variable as well as the contributions to the output of a response function, since the response function is a deterministic transfer function from the inputs $\boldsymbol{X}$ to the output Y. The propagation of the results of Bayesian model averaging through a response function will be discussed later in Section 3.6.3.

3.6.1.4 Summary

Conventionally, model averaging methods assign weights for competing models, and these weights are estimated in a deterministic manner. In this section, the uncertainty in the weights is also computed, thereby giving both the confidence in a particular distribution type (through the mean value of w) , and a measure of uncertainty in this confidence (through the standard deviation of w). One disadvantage of this approach is that it assumes spurious interactions between competing model forms while constructing the joint likelihood of weights and distribution parameters of all model forms. As a result, this approach involves multi-dimensional integration; a significant amount of computational power may be required, if there are several competing model forms. For example, if there were 5 competing model forms, each with two distribution parameters, then the joint likelihood needs to be constructed for 14 quantities (4 weights and 10 parameters), and a 14-dimensional integration is needed to quantify the model form uncertainty and estimate the distribution parameters.

The next section discusses the use of Bayesian hypothesis testing to quantify model form uncertainty; this approach provides a computationally efficient alternative and also directly computes the probability that the data supports a given model form.

3.6.2 Bayesian Hypothesis Testing Approach

Similar to the previous section, consider the two PDFs $f_X^1(x|\boldsymbol{\phi})$ and $f_X^2(x|\boldsymbol{\theta})$ to be the two competing model forms M_1 and M_2 respectively. In Bayesian hypothesis testing, prior probabilities ($P(M_1)$ and $P(M_2)$) are assumed for each of these events, and Bayes theorem is used to update their probabilities based on the available data (D), as [91]:

$$\frac{P(M_1|D)}{P(M_2|D)} = \frac{P(D|M_1)}{P(D|M_2)} \frac{P(M_1)}{P(M_2)} \tag{3.19}$$

The first term on the right hand side of Eq. 3.19 is referred to as the Bayes factor, denoted by B [17].

$$B = \frac{P(D|M_1)}{P(D|M_2)} \tag{3.20}$$

The Bayes factor is the ratio of likelihoods of M_1 and M_2 and is a quantitative measure of extent of data support for model M_1 relative to the support for M_2. If $B > 1$, then the data D favors model M_1. Higher the Bayes factor, higher is the likelihood of the model M_1. In the absence of any prior preference between M_1 and M_2, assume equal prior probabilities, i.e. $P(M_1) = P(M_2) = 0.5$. Then, the posterior probabilities ($P(M_1|D)$ and $P(M_2|D)$) can be expressed in terms of the Bayes factor as:

$$\begin{aligned} P(M_1|D) &= \frac{B}{B+1} \\ P(M_2|D) &= \frac{1}{B+1} \end{aligned} \tag{3.21}$$

In order to implement this, the likelihood functions ($P(D|M_1)$ and $P(D|M_2)$) must be calculated. This is accomplished in two steps. In the first step, $P(D|M_1, \boldsymbol{\phi})$ is calculated using the data D available. Similar to the Section 3.6.1, assume that m point data x_i ($i = 1$ to m) and n intervals $[a_i, b_i]$ ($i = 1$ to n) are available.

$$P(D|M_1, \boldsymbol{\phi}) \propto L(M_1, \boldsymbol{\phi}) = \prod_{i=1}^{m} f_X^1(x = x_i|\boldsymbol{\phi}) \prod_{j=1}^{n} \int_{a_j}^{b_j} f_X^1(x|\boldsymbol{\phi}) dx \tag{3.22}$$

Similarly, $P(D|M_2, \boldsymbol{\theta})$ is also calculated. In the second step, these two quantities are used to calculate $P(D|M_1)$ and $P(D|M_2)$. Let $f_{\boldsymbol{\phi}}(\boldsymbol{\phi})$ denote the prior PDF of the distribution parameter $\boldsymbol{\phi}$. Using conditional probability, it follows that

$$L(M_1) \propto P(D|M_1) = \int P(D|M_1, \boldsymbol{\phi}) f_{\boldsymbol{\phi}}(\boldsymbol{\phi}) d\boldsymbol{\phi} \tag{3.23}$$

If a uniform prior density is assigned for $\boldsymbol{\phi}$, then the above equation reduces to

$$L(M_1) \propto P(D|M_1) \propto \int P(D|M_1, \boldsymbol{\phi}) d\boldsymbol{\phi} \tag{3.24}$$

Using Eq. 3.21, the posterior probability of model M_1, i.e. $P(M_1|D)$ can be calculated. Similar equations can be written for model M_2.

The evaluation of the above probabilities involves multi-dimensional integration; however the number of dimensions is only equal to the number of distribution parameters for each individual distribution. In contrast, the Bayesian model averaging approach discussed earlier in Section 3.6.1 would require multi-dimensional integration with all weights and parameters together. Hence, Bayesian hypothesis testing is computationally more affordable in comparison with the Bayesian model averaging approach.

3.6.2.1 Single and Multiple Model Forms

The case of two competing models was discussed above. This method can be extended to (1) addressing model form uncertainty in a single model; and (2) quantifying the model form uncertainty for multiple models.

Consider the case is when there is only one model M_1 and it is desired to calculate the model form uncertainty. This can be viewed as a hypothesis testing problem where the null hypothesis is that model M_1 is correct, and alternate hypothesis is that model M_2 is correct, where model M_2 is the opposite of model M_1. One possible approach is to choose the model M_2 as a uniform distribution (non-informative). Hence, $f_X^2(x|\boldsymbol{\theta})$ is a uniform PDF; the PDFs of the lower and upper bounds are estimated based on the data and then "integrated out" to compute $P(M_1|D)$ and $P(M_2|D)$.

If there are more than two competing models, say n models, then the Bayes factor which was earlier defined as a ratio between two models can now be defined in

terms of proportions as $P(D|M_1) : P(D|M_2) : P(D|M_3)...P(D|M_n)$. Using equations analogous to those in the previous subsection, the probabilities $P(M_1|D)$, $P(M_2|D)$, $P(M_3|D)$ and so on until $P(M_n|D)$ can also be calculated.

The following subsections present two illustrations to show how the proposed methodology works. These examples are the same as those in Section 3.6.1, and used to illustrate the usage of Bayesian hypothesis testing for quantifying the distribution type uncertainty.

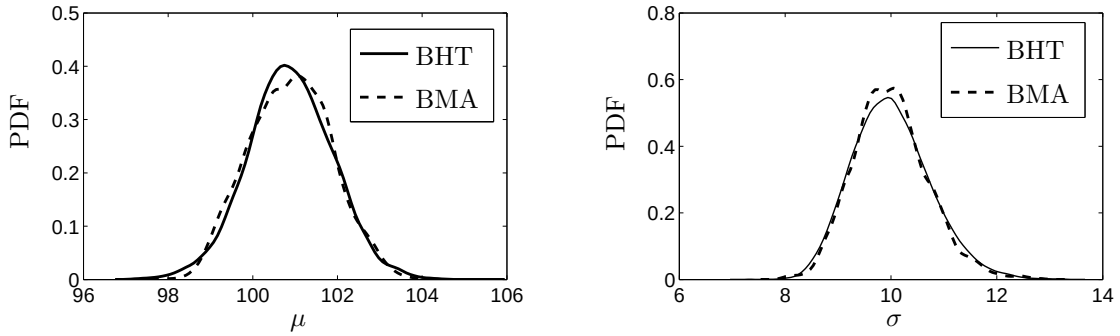
3.6.2.2 Illustration 1

Consider the same data set as in Section 3.6.1.1, i.e. 100 samples drawn from N(100,10). The two competing model forms are normal ($M_1 : N(\mu, \sigma)$) and uniform ($M_2 : U(a, b)$).

Using the Bayes factor, the probabilities $P(M_1)$ and $P(M_2)$ are found to be one and zero (upto 5th decimal place), thereby isolating the normal distribution with almost 100% confidence. This behavior is similar to that in the Bayesian model averaging method. The reasons for this behavior are the same as those previously mentioned: (1) sufficient data to uniquely identify the normal distribution; (2) significant difference between the two competing model forms, normal and uniform.

Similar to the Bayesian model averaging (BMA) procedure, the PDFs of the distribution parameters using the Bayesian hypothesis testing (BHT) approach are also quantified, and shown in Fig. 3.17. Note that the results from both the approaches are shown for the sake of comparison. Since the normal distribution has been isolated with almost 100% confidence, the distribution parameters are shown only for normal distribution.

Note that there is no significant difference between the PDFs of the distribution parameters estimated through the Bayesian hypothesis testing route or the model



(a) PDF of μ (b) PDF of σ

Figure 3.17: PDFs of Distribution Parameters

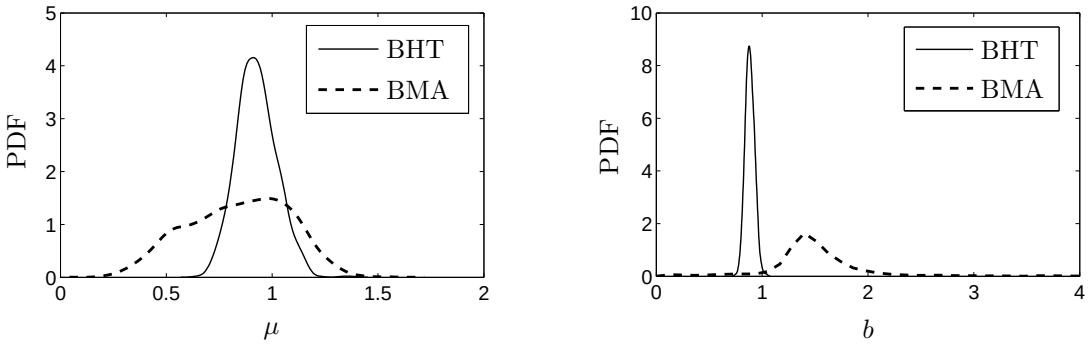
averaging route. This is expected because both the methods completely isolate the model form to normal distribution (which means BMA did not impose interactions between distribution parameters of the two competing model forms), and hence, the PDFs of the distribution parameters are expected to be the same. The difference between the two methods is only in the quantification of the model form uncertainty, and the computational effort.

3.6.2.3 Illustration 2

Consider 100 samples of data generated from an exponential distribution. The two competing model forms M_1 and M_2 are exponential and Rayleigh distributions respectively. The probabilities $P(M_1|D)$ and $P(M_2|D)$ are also estimated using Eq. 3.21. These posterior probabilities are found to be 1 and 0 respectively.

Further, the PDFs of the distribution parameters, i.e. μ for the exponential distribution and b for the Rayleigh distribution are also quantified, and shown in Fig. 3.18. Similar to the previous numerical example, the results from the Bayesian model averaging approach are also provided for the sake of comparison.

There are two important observations. First, these results are considerably different from the Bayesian model averaging results. Second, the uncertainty (measured in



(a) PDF of μ (b) PDF of b

Figure 3.18: PDFs of Distribution Parameters

terms of standard deviation) in the results from Bayesian hypothesis testing is much smaller than that from the model averaging approach.

This behavior is due to the conceptual differences between the two approaches. The Bayesian model averaging approach considers the joint likelihood of weights and parameters of all distribution types, thereby assuming interactions between all the parameters (where there is none). In contrast, the hypothesis testing approach only considers the joint likelihood of all parameters of a single distribution type and does not include interactions across multiple distribution types. As a result, the estimation of μ in the hypothesis testing approach is completely independent of b; on the contrary, these two parameters are estimated simultaneously in the model averaging approach. The results of Bayesian hypothesis testing have smaller uncertainty because fewer parameters are estimated with the same amount of data.

These differences were not seen in the first numerical example because the normal distribution and the uniform distribution are significantly different from each other and the data wholly supported the normal distribution; whereas the exponential and Rayleigh distributions are not.

3.6.2.4 Quantifying Individual Contributions

Now, the goal is to quantify the individual contributions of (1) variability; (2) distribution type uncertainty; and (3) distribution parameter uncertainty. Earlier, in the Bayesian model averaging approach (Section 3.6.1.3), the distribution type uncertainty was represented by a continuous random variable (w). On the other hand, now the difference is that the distribution type uncertainty is represented using a discrete random variable.

Without loss of generality, consider two distribution types (M_1 and M_2) and the corresponding probabilities ($P(M_1)$ and $P(M_2)$) estimated using the Bayesian hypothesis testing method. First, a discrete random number (denoted by T and uniformly distributed on [0, 1]) needs to be sampled based on the value of $P(M_1|D)$ to select between the competing models, i.e. M_1 and M_2. Based on the sampled a value of T, the distribution type is selected. Given a value of distribution parameter, then X is represented using a PDF. Now, an auxiliary variable U_X is defined as:

$$\begin{aligned} U_X &= \int_{-\infty}^{X} f_X^1(x|\boldsymbol{\phi})dx \text{ if } T < P(M_1) \\ U_X &= \int_{-\infty}^{X} f_X^2(x|\boldsymbol{\theta})dx \text{ if } T > P(M_1) \end{aligned} \tag{3.25}$$

Similar to the Sections 3.5 and 3.6.1.3, U_X is uniformly distributed on [0, 1] and the above equations provide a deterministic model to carry out global sensitivity analysis. The contribution of physical variability is calculated as:

$$\begin{aligned} S_P^I &= \frac{V_{U_X}(E_{T,\boldsymbol{\phi},\boldsymbol{\theta}}(X|U_X))}{V(X}) \\ S_P^O &= 1 - \frac{V_{T,\boldsymbol{\phi},\boldsymbol{\theta}}(E_{U_X}(X|T,\boldsymbol{\phi},\boldsymbol{\theta}))}{V(X)} \end{aligned} \tag{3.26}$$

Table 3.7: Contributions of Physical Variability and Epistemic Uncertainty

Illustration	Effect	Physical Variability	Epistemic Uncertainty	Distribution Type
Example 1	Individual	94.0%	4.0%	0.0%
Section 3.6.2.2	Overall	98.0%	5.0%	0.0%
Example 2	Individual	72.7%	25.4%	0.0%
Section 3.6.2.3	Overall	75.5%	30.3%	0.0%

In Eq. 3.26, S_P^I and S_P^O represent the individual effect and overall effect of physical variability respectively.

Since the distribution parameter is calculated only after selecting the distribution type, it is not meaningful to calculate the contribution of distribution parameter uncertainty alone. The individual and total effects of epistemic uncertainty (i.e. distribution parameter uncertainty and distribution type uncertainty) can be calculated as:

$$\begin{aligned} S_E^I &= \frac{V_{T,\boldsymbol{\phi},\boldsymbol{\theta}}(E_{U_X}(X|T,\boldsymbol{\phi},\boldsymbol{\theta}))}{V(X}) \\ S_E^O &= 1 - \frac{V_{U_X}(E_{T,\boldsymbol{\phi},\boldsymbol{\theta}}(X|U_X))}{V(X)} \end{aligned} \tag{3.27}$$

Also, the individual and total effects of distribution type uncertainty can be calculated as:

$$\begin{aligned} S_{Type}^I &= \frac{V_T(E_{U_X,\boldsymbol{\phi},\boldsymbol{\theta}}(X|T))}{V(X}) \\ S_{Type}^O &= 1 - \frac{V_{U_X,\boldsymbol{\phi},\boldsymbol{\theta}}(E_T(X|U_X,\boldsymbol{\phi},\boldsymbol{\theta}))}{V(X)} \end{aligned} \tag{3.28}$$

The above equations calculate the contributions in a single variable. For example, the results of sensitivity analysis for the illustrative examples (discussed earlier in Sections 3.6.2.2 and 3.6.2.3) are tabulated in Table 3.7. Note that the individual and

total contributions of distribution type uncertainty are zero in both the examples, because it was possible to isolate one distribution type uniquely in both the examples.

Similar to Section 3.5, it is straightforward to extend the above equations to quantify the individual contributions to the output of a response function, since the response function is a one-to-one mapping between the inputs and the output. The propagation of the results of Bayesian hypothesis testing through a response function will be discussed later in Section 3.6.3.

3.6.2.5 Summary

The Bayesian hypothesis testing approach quantifies the distribution type uncertainty through the posterior probability ($P(M_1|D)$) which is deterministic in contrast with the model averaging approach which calculates a stochastic weight (w). It is clear that the Bayesian model averaging and Bayesian hypothesis testing methods are based on different assumptions; they are conceptually different and caution must be exercised while comparing the results of these methods. From the perspective of computational efficiency, it may be advantageous to use Bayesian hypothesis testing, thereby not allowing spurious interactions between distribution parameters of multiple model forms.

The Bayesian hypothesis testing method can also be used when the PDFs of the distribution parameters of two competing model forms are readily available. For each realization of distribution parameter values, the Bayes factor is calculated, thereby leading to the PDF of the Bayes factor [92]. This approach is significantly different from the concern in this chapter, where the probability that the model is correct and the PDF of the corresponding distribution parameters are estimated *simultaneously* using the available data, thereby leading to a single Bayes factor value which is easier for the purpose of decision making.

3.6.3 Uncertainty Propagation through a Model

Consider the case where the quantity X is an input to a mathematical model ($Y = g(\boldsymbol{X})$,), and all three types of uncertainty - physical variability, distribution type and distribution parameters - in X need to propagated through the system model to compute the uncertainty in the response Y. This section discusses the various issues in such uncertainty propagation and numerical implementation of the uncertainty propagation.

3.6.3.1 Propagation using Bayesian Model Averaging

In the Bayesian model averaging approach, the PDFs of w, $\boldsymbol{\phi}$, and $\boldsymbol{\theta}$ are all calculated simultaneously. In other words, the joint likelihood of these quantities is used to calculate the individual marginal PDFs. Based on Eq. 3.14, a given realization of w, $\boldsymbol{\phi}$, and $\boldsymbol{\theta}$ values lead to a particular PDF $f_X(x|w, \boldsymbol{\phi}, \boldsymbol{\theta})$. Let $F_X(x|w, \boldsymbol{\phi}, \boldsymbol{\theta})$ denote the corresponding CDF. For multiple values of w, $\boldsymbol{\phi}$, and $\boldsymbol{\theta}$. there exists a family of PDFs for X. Each PDF can be propagated through the above response function, and a family of PDFs for Y can be calculated.

However, in practice, it may not be possible to directly invert the CDF $F_X(x|w, \boldsymbol{\phi}, \boldsymbol{\theta})$; thus a composite method is used [88]. For a set of sampled values of w, $\boldsymbol{\phi}$, and $\boldsymbol{\theta}$, this CDF can be inverted numerically. A uniform random number on [0, 1] is drawn. If this number is less than the sampled value of w, then a random sample of X is selected from the PDF $f_X^1(x|\boldsymbol{\phi})$; else a random sample of X is drawn from the PDF $f_X^2(x|\boldsymbol{\theta})$. Multiple such samples of X correspond to the PDF $f_X(x|w, \boldsymbol{\phi}, \boldsymbol{\theta})$. This procedure is repeated for multiple samples of w, $\boldsymbol{\phi}$, and $\boldsymbol{\theta}$ to generate a family of distributions for X. Note that this procedure is different from the algorithm in Fig. 3.11; here, both the distribution type and distribution parameters are sampled at the same level, whereas in Fig. 3.11, they were sampled one within the other.

For the purpose of uncertainty propagation, the family of distributions approach may be computationally expensive because it needs two Monte Carlo loops, one within the other. In that case, the family of distributions may be replaced with a single, unconditional PDF $f_X(x)$ by integrating over w, $\boldsymbol{\phi}$, and $\boldsymbol{\theta}$, as:

$$f_X(x) = \int f_X(x|w, \boldsymbol{\phi}, \boldsymbol{\theta}) f(w|D) f(\boldsymbol{\phi}|D) f(\boldsymbol{\theta}|D) dw d\boldsymbol{\phi} d\boldsymbol{\theta} \tag{3.29}$$

The above integral can be numerically evaluated using sampling. In the aforementioned double-loop sampling procedure, several values of X were chosen for a given sample of w, $\boldsymbol{\phi}$, and $\boldsymbol{\theta}$, thereby establishing conditional dependence on w, $\boldsymbol{\phi}$, and $\boldsymbol{\theta}$ and leading to the PDF $f_X(x|w, \boldsymbol{\phi}, \boldsymbol{\theta})$. In order to compute the single, unconditional PDF as in Eq. 3.29, only one sample of X is chosen for a given sample of w, $\boldsymbol{\phi}$, and $\boldsymbol{\theta}$. This sampling procedure may be referred to as single-loop sampling. The resultant unconditional, predictive PDF includes all three types of uncertainty - variability, distribution type and distribution parameter - and can be used for uncertainty propagation through the system model $Y = g(X)$.

It may appear that the use of the unconditional, predictive PDF may lead to loss of information regarding the individual contributions of the three aforementioned types of uncertainty. It may be desirable to assess their individual contributions. Currently, this issue has been addressed only qualitatively through graphical visualization, as seen earlier in Fig. 3.12. Future research needs to address the rigorous quantification of the individual contributions of the three types of uncertainty.

3.6.3.2 Propagation using Bayesian Hypothesis Testing

While the Bayesian model averaging approach leads to a stochastic weight w, the Bayesian hypothesis testing approach leads to a deterministic posterior probability $P(M_1|D)$. Hence, obtaining the family of distributions is simpler than in the Bayesian

model averaging approach because the uncertainty in the weight is not considered. However, the order of sampling is different. First, a uniform random number on [0, 1] is drawn. If this uniform random number is less than $P(M_1|D)$, then a random sample of $\boldsymbol{\phi}$ is selected and multiple samples of X are drawn from the PDF $f_X^1(x|\boldsymbol{\phi})$; this procedure is repeated for several samples of $\boldsymbol{\phi}$. If the uniform random number is greater than $P(M_1|D)$, then a random sample of $\boldsymbol{\theta}$ is selected and multiple samples of X are drawn from the PDF $f_X^2(x|\boldsymbol{\theta})$; this procedure is repeated for several samples of $\boldsymbol{\theta}$. This algorithm is, in fact, exactly the same as that in Fig. 3.11 and leads to a family of PDFs similar to that in Fig. 3.12.

As stated earlier, the family of PDFs approach is computationally expensive for the purpose of uncertainty propagation. Hence, a simultaneous sampling approach is used to construct a single, unconditional PDF of X. In this simultaneous sampling approach, given a uniform random number sample, the distribution type is selected, and only one sample of distribution parameters (of the selected distribution type), and only one sample of X are selected. The procedure is repeated for multiple uniform random number samples, and multiple samples of X are obtained, thereby leading to the unconditional PDF of X. Similar to that in Bayesian model averaging, this unconditional, predictive PDF can be used for the purpose of uncertainty propagation through the system model $Y = g(X)$.

3.7 Case 3: Unknown PDF Type (Non-parametric)

The need to quantify the uncertainty due to the choice of a particular distribution type was emphasized in Section 3.4.5, where it was demonstrated that different assumptions regarding the distribution type lead to different unconditional PDFs. One method to address distribution type uncertainty was using multiple competing distribution types, as demonstrated in Section 3.6. This section proposes another

method where no distribution type is assumed and a non-parametric distribution is constructed which faithfully represents the available data. As a result, there is no distinction between variability and distribution parameter uncertainty in the resultant non-parametric distribution; from a subjective probability-based point of view, it simply represents the degree of belief regarding the quantity of interest.

The non-parametric approach is based on the fact that if the PDF values are known at a few points, then the entire density function can be constructed based on an interpolation method. No explicit PDF form such as $f_X(x|\boldsymbol{P})$ is assumed, and the interpolation method is not parametric. Since there are no distribution parameters or interpolation parameters, the method is referred to as being non-parametric.

Discretize the domain of the input quantity X into a finite number of points, say λ_i, $i = 1$ to Q. This domain is chosen based on available data; the lowest value and the highest value are chosen as the lower bound and the upper bound of the domain respectively. Assume that the PDF values, i.e. $f_X(x = \lambda_i)$ at each of these Q points are given by $f_X(x = \lambda_i) = \eta_i$ for $i = 1$ to Q. Using an interpolation technique, the entire probability density function $f_X(x)$ can be calculated for all $\lambda \in X$, i.e. over the entire domain of X. Then the probability of observing the given data (point data and interval data), i.e. the likelihood, can be calculated using Eq. 3.6. This likelihood is a function of the following:

1. The discretization points selected, i.e. λ_i, $i = 1$ to Q;

2. The corresponding PDF values $f_X(x = \lambda_i) = \eta_i$, $i = 1$ to Q; and

3. The type of interpolation technique used.

In this research, the discretization is fixed, i.e. uniformly spaced λ_i values ($i = 1$ to Q) over the domain of X are chosen in advance and the values of η_i that maximize the likelihood function are calculated. The value of Q (number of discretization points) is

chosen based on computational power; higher the value of Q, better the results, and more expensive the optimization. The optimization problem is formulated as shown in Eq. 3.30.

$$\begin{gathered}
\text{Given } \lambda_i \in X \ \forall \ i, i = 1 \text{ to } Q \\
\underset{\boldsymbol{\eta}}{\text{Maximize}} \ L(\boldsymbol{\eta}) \\
\boldsymbol{\eta} = \{\eta_1, \eta_2, ...\eta_{Q-1}, \eta_Q\} \text{ and } f_X(x = \lambda_i) = \eta_i \ \forall \ i = 1 \text{ to } Q \\
\text{subject to} \\
\eta_i \geq 0 \ \forall \ i = 1 \text{ to } Q \\
f_X(x) \geq 0 \text{ and } \int f_X(x) = 1 \\
\text{PDF values } \eta_i \text{ at } x = \lambda_i \text{ used to interpolate entire PDF } f_X(x) \ \forall \ x
\end{gathered} \tag{3.30}$$

The objective of this optimization is to maximize the likelihood function $L(\boldsymbol{\eta})$ (calculated using the point and interval data, similar to that in Eq. 3.6), now the difference being that there are no distribution parameters here. The optimization is performed subject to the following constraints: the vector $\boldsymbol{\eta}$ contains PDF values that need to be positive; the resultant function $f_X(x)$ must satisfy the properties of a probability density function, i.e. it must be positive and integrate to unity.

Different interpolation techniques – linear interpolation, cubic spline-based interpolation and the Gaussian process (GP) interpolation - are investigated. While the method of linear interpolation is based on piecewise linear approximation, the method of spline interpolation minimizes the integral of the squared curvature for approximation [93]. The method of Gaussian process interpolation was explained earlier in Section 2.8. The resulting non-parametric PDFs from the three interpolation methods are plotted in Fig. 3.19. From Fig. 3.19, it can be seen that the method constructs the PDF that closely follows the available data. Because there was no data available in the range 4.5 to 5.0, the method assigns a low relative likelihood to this range of

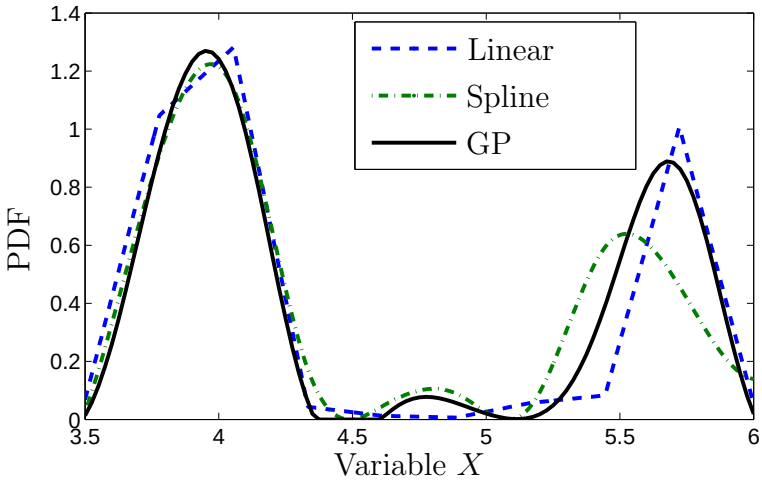


Figure 3.19: Non-parametric Probability Distributions

values. Further, the method also allows the appearance of multiple peaks leading to a multi-modal density function; the density function assumes this shape depending on the available data.

While the linear and spline interpolation techniques use linear and cubic functions for interpolation, the Gaussian process interpolation method does not assume any explicit functional form for the interpolation. Further, it was mentioned earlier that non-parametric methods do not make assumptions about the distribution form. Hence, the Gaussian process-based approach can be described as the most general approach; it neither assumes a functional form for the PDF nor uses explicit polynomial functions for the interpolation technique. Since the non-parametric approach directly calculates a single distribution for the quantity of interest, it is not possible to assess the individual contributions of variability and epistemic uncertainty.

The above non-parametric technique is not useful when only one interval is available; it can be useful when sparse point and/or multiple interval data are available

on a particular quantity. This quantity may be the random variable itself or a distribution parameter of the random variable. While the above numerical example considered the former case, the latter case is considered in Section 3.8.

3.8 Sandia Challenge Problem

This section uses the proposed methods to solve the challenge problems considered in the Sandia Epistemic Uncertainty Workshop [57]. In these problems, at least one input quantity is described using intervals and the uncertainty in model output needs to be quantified. Using the proposed methodology, the inputs can be modeled as PDFs which are then used for uncertainty propagation. Since the focus of this chapter is on uncertainty representation and quantification rather than uncertainty propagation, a simple Monte Carlo sampling-based method is used for uncertainty propagation instead of advanced sampling techniques or analytical approximations such as the First Order Reliability Method (FORM) and the Second Order Reliability Method (SORM). Any of these methods can be used instead of the simple sampling approach here.

The aforementioned uncertainty workshop originally consists of two sets of Challenge problems - A and B. These two sets consist of several uncertainty propagation problems based on the type of data available on the inputs. This section considers only two problems (3^{rd} and 5^{th}) from Set A where multiple interval data are available for the same quantity. The rest of the problems have single interval description; it may be convenient to assume a uniform distribution based on the principle of maximum entropy in such cases or simply perform interval analysis using the single interval. The real challenge lies when multiple interval data are available for the same quantity. These multiple intervals are assumed to be statistically independent so that the likelihood-based approach can be implemented. A hypothetical model with two

independent inputs a and b is considered as:

$$y = (a + b)^a \tag{3.31}$$

The two uncertainty propagation problems are discussed in detail, in the following subsections.

3.8.1 Problem 1

The uncertainty in the input variable a is described using three intervals [0.5, 0.7], [0.3, 0.8] and [0.1, 1.0]. Input variable b follows a lognormal distribution, but with imprecise distribution parameters (λ and ξ). These distribution parameters are described by multiple intervals ([0.6, 0.8], [0.2, 0.9], [0.0, 1.0]) and ([0.3, 0.4], [0.2, 0.45], [0.1, 0.5]), respectively. In this problem:

1. The distribution parameters (λ and ξ) of b are described with intervals. It may not be meaningful to represent λ and ξ using parametric distributions because this approach would lead to distribution parameters of distribution parameters. Hence, the Gaussian process-based non-parametric distribution is used to represent λ and ξ. For the sake of uniformity, a is also represented using a similar non-parametric distribution. So, there is not issue of variability versus distribution parameter uncertainty for the input a.

2. The variable b is explicitly given to follow a lognormal distribution, and hence using this problem, it is possible to apportion the uncertainty in y to (1) uncertainty in a, given by the non-parametric PDF; (2) lognormal variability in b; and (3) uncertainty in distribution parameters (λ and ξ) of b.

The PDFs of a, λ, and ξ are first computed using the non-parametric method described in Section 3.7. These PDFs are plotted in Figs 3.20 and 3.21.

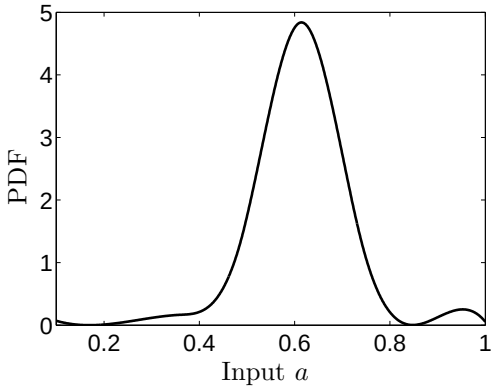
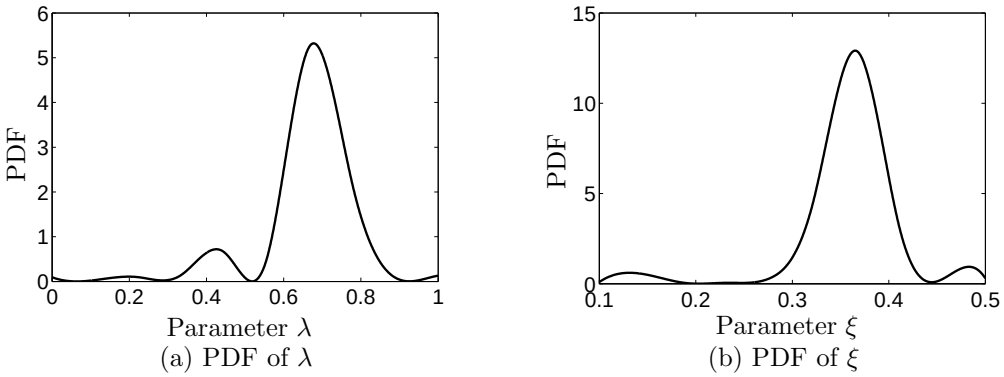


Figure 3.20: PDF of Input a



(a) PDF of λ (b) PDF of ξ

Figure 3.21: PDFs of Distribution Parameters of Input b

Every realization of λ and ξ leads to a PDF of b, and hence to a PDF of output y. Since λ and ξ are themselves uncertain, y is represented using a family of PDFs. Alternative it is also possible to compute the unconditional PDF of b, as:

$$f(b) = \int f(b|\lambda, \xi) f(\lambda) f(\xi) d\lambda d\xi \tag{3.32}$$

This unconditional PDF of b is used to compute the unconditional PDF of y, which is plotted in Fig. 3.22. The proposed methodology produces a complete PDF for y, some of the previously existing solution methodologies result in interval values. Kozine and Utkin [94], De Cooman and Troffaes [95], and Ferson and Hajagos [96]

produced [1.5, 2.8], [1.5, 2.2], and [1.1, 3.8] respectively. Zaman et al. [61] produced a family of distributions; while it is challenging to compare a family of PDFs against the proposed single PDF, a qualitative comparison shows considerable agreement between the two approaches.

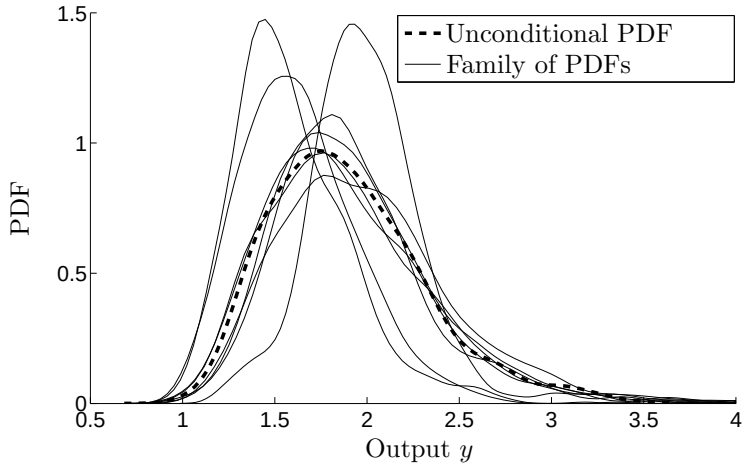


Figure 3.22: Output y: Family of PDFs and Unconditional PDF

The above exercise sets the problem ready for sensitivity analysis. The uncertainty in y can be attributed to: (1) uncertainty in a, shown in Fig. 3.20; (2) variability in b, which is a lognormally distributed quantity; as per the proposed sensitivity analysis method, an auxiliary variable U is introduced to represent this variability; and (3) uncertainty in distribution parameters (λ and ξ), shown in Fig. 3.21. There are 15 terms in the decomposition of variance.

Table 3.8: Contributions of Variability and Distribution Parameter Uncertainty

Quantity	**Meaning**	**Individual Effects**	**Overall Effects**
(λ, ξ)	Parameter Uncertainty in b	8.0 %	10.4 %
U	Variability in b	56.8 %	62.5 %
(λ, ξ, U)	Total uncertainty in b	66.7 %	70.9 %
a	Uncertainty in a	29.1 %	33.3 %

Table 3.8 reports the individual and overall sensitivity indices (to one decimal place) of the following quantities: (1) parameter uncertainty in b; (2) variability in b; (3) total uncertainty in b; and (4) uncertainty in a. It is seen that there is little interaction between variability and parameter uncertainty of b. The contribution of distribution parameter uncertainty is about 10%, and the variability in b contributes to about 60% of the overall variance; while the former can be reduced, the latter is irreducible uncertainty.

3.8.2 Problem 2

In this problem, both the input variables a is described using three intervals ([0.5, 0.7], [0.3, 0.8], [0.1, 1.0]) and the input variable b is described using three intervals ([0.4, 0.85], [0.2, 0.9], [0.0, 1.0]) and one point datum (0.6). In this problem, intervals are available for the quantities directly, as against intervals for distribution parameters. Hence, it is meaningful to pursue a parametric approach and compute the distribution parameters of both a and b. For the sake of illustration, the two competing parametric distribution types - normal and uniform - are chosen for both a and b. The goal is to quantify the distribution type uncertainty in a and b, the corresponding distribution parameters, and then propagate the three types of uncertainty - variability, distribution type uncertainty, and distribution parameter uncertainty - to compute the uncertainty in y. The methods of Bayesian model averaging and Bayesian hypothesis testing are used for this purpose.

3.8.2.1 Bayesian Model Averaging

For the quantity a, model M_1 is chosen as $\mathrm{N}(\mu_a,\sigma_a)$, and model M_2 is chosen as $\mathrm{U}(L_a,U_a)$. Similarly, for the quantity b, model M_1 is chosen as $\mathrm{N}(\mu_b,\sigma_b)$, and model M_2 is chosen as $\mathrm{U}(L_b,U_b)$. Let w_a and w_b denote the weights assigned to the normal

distribution for the variables a and b respectively. Then, $1 - w_a$ and $1 - w_b$ denote the weights assigned to the uniform distribution for the variables a and b respectively. For the sake of illustration, the prior distribution for all the distribution parameters is chosen to be uniform on the interval [0, 1].

The PDFs of all the above quantities are estimated using the available data, and the mean values, the standard deviations, maximum likelihood estimates, and 95% bounds are shown in Table 3.9.

Table 3.9: Bayesian Model Averaging: Results

Variable	Distribution Type	Quantity	Mean	Standard Deviation	95% Bounds
a	Normal	w_a	0.36	0.23	[0.04,0.76]
		μ_a	0.50	0.30	[0.03, 0.93]
		σ_a	0.20	0.25	[0.03, 0.75]
	Uniform	$1 - w_a$	0.64	0.23	[0.01,0.96]
		L_a	0.73	0.17	[0.42, 0.91]
		U_a	0.86	0.13	[0.61, 0.99]
b	Normal	w_b	0.58	0.29	[0.05, 0.97]
		μ_b	0.57	0.15	[0.23, 0.84]
		σ_b	0.35	0.53	[0.01, 1.34]
	Uniform	$1 - w_b$	0.42	0.29	[0.03, 0.95]
		L_b	0.37	0.19	[0.04, 0.62]
		U_b	0.67	0.25	[0.29, 0.95]

The results in Table 3.9 are difficult to interpret for a number of reasons, the primary reason being that all the estimates have very high degree of uncertainty (indicated by standard deviation). This happens because the method tries to estimate 5 parameters simultaneously using a small data set (3 for a and 4 for b). As a result, the 95% bounds are too large to be useful. The PDFs of the weights w_a and w_b are almost uniform, suggesting that even the maximum likelihood estimates may not be useful. Also, consider the uniform distribution estimated for a; the estimates of the lower and higher bounds are so close (but with high standard deviations) that it is difficult to derive any usefulness from such results.

Table 3.10: Bayesian Hypothesis Testing Results

Variable	$P(M\|D)$	Quantity	Mean	Standard Deviation	95% Bounds
a	Normal	μ_a	0.57	0.16	[0.16,0.89]
	0.32	σ_a	0.23	0.20	[0.01, 0.80]
	Uniform	L_a	0.41	0.16	[0.05, 0.65]
	0.68	U_a	0.74	0.12	[0.54,0.97]
b	Normal	μ_b	0.60	0.12	[0.30, 0.89]
	0.28	σ_b	0.17	0.15	[0.01, 0.64]
	Uniform	L_b	0.43	0.15	[0.08, 0.60]
	0.72	U_b	0.74	0.11	[0.60, 0.97]

Due to the large uncertainty in the input, further uncertainty propagation analysis is not useful. Instead, the Bayesian hypothesis testing approach is investigated next. Note that the hypothesis testing approach does not estimate more than 2 parameters simultaneously, and hence is expected to produce results that have less uncertainty.

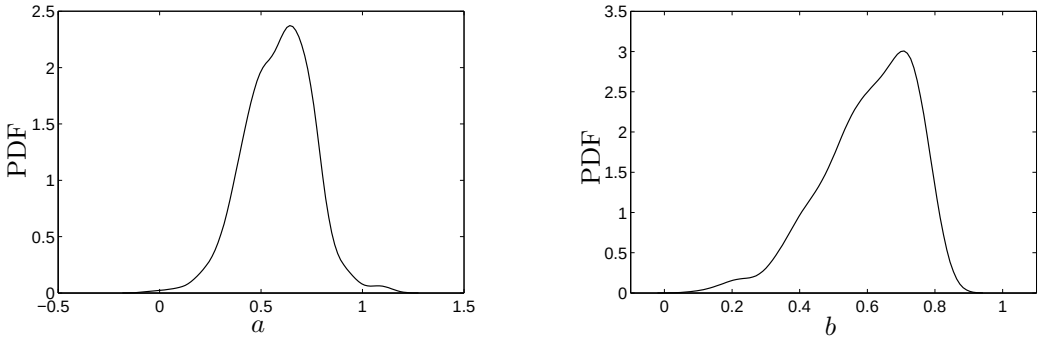
3.8.2.2 Bayesian Hypothesis Testing

Using the Bayesian hypothesis testing approach proposed in Section 3.6.2, the probabilities $P(M_1|D)$ and $P(M_2|D)$ can be directly calculated for both a and b. Then, the PDFs of the distribution parameters (μ_a and σ_a for normal, and L_a and U_a for uniform) can also be calculated. The results of the distribution parameter estimation are shown in Table 3.10. Note that the estimation of the parameters of the normal distribution is totally independent of the estimation of the parameters of the uniform distribution, for both the variables a and b. However, this was not the case in the Bayesian model averaging approach.

Once the uncertainty in the model form and the distribution parameters are estimated, then a and b are represented using a single, unconditional PDF each. The calculation of this unconditional PDF is using the simultaneous sampling approach

explained earlier in Section 3.6.3.2. This single PDF accounts for physical variability, distribution type uncertainty and distribution parameter uncertainty, and hence renders the uncertainty propagation analysis efficient.

The unconditional PDF of a and the unconditional PDF of b is then propagated through Eq. 3.31 to calculate the PDF of Y. The PDFs of a and b are shown in Fig. 3.23.



(a) PDF of a (b) PDF of b

Figure 3.23: PDFs of Model Inputs

Using simple Monte Carlo simulation, the PDF of y is then calculated and shown in Fig 3.24. Since this PDF is calculated from multiple competing parametric PDFs, it is indicated as the parametric PDF. It accounts for all sources of uncertainty in the inputs - physical variability, distribution type uncertainty and distribution parameter uncertainty.

The computational method developed in Section 3.6.2.4 is then used to quantify the individual contributions of aleatory (physical variability) and epistemic uncertainty (distribution type and parameter uncertainty). Recall that it is not meaningful to calculate the effect of distribution parameter uncertainty alone, because the choice of distribution parameters is made only after selecting the distribution type. The individual and overall effects for all the three quantities a, b, and y are tabulated in Table 3.11.

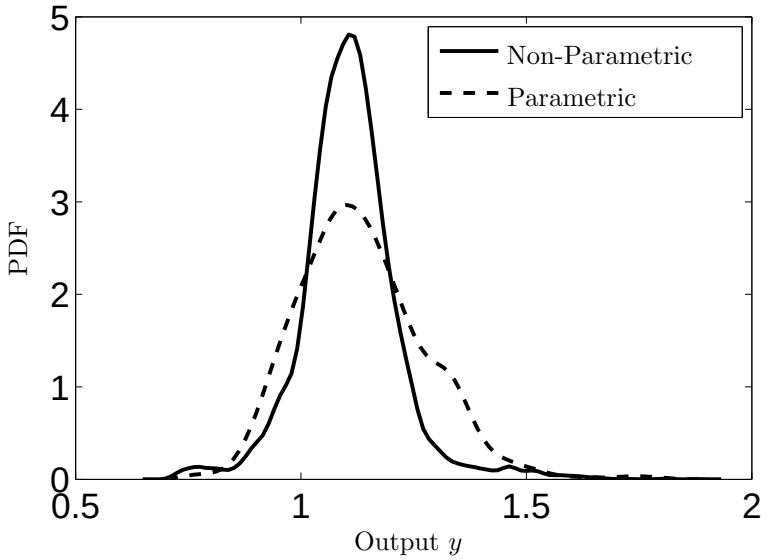


Figure 3.24: PDF of Model Output y

Table 3.11: Contributions of Physical Variability and Epistemic Uncertainty

Quantity	Effect	Physical Variability	Epistemic Uncertainty	Distribution Type
a	Individual	31.5%	20.8%	0.0%
	Overall	79.2%	68.5%	43.3%
b	Individual	23.7%	36.5%	0.0%
	Overall	63.5%	77.3%	9.9%
y	Individual	4.2%	4.1%	1.0%
	Overall	95.9%	95.8%	56.0%

It is seen from the results that the individual effect of distribution type uncertainty is almost zero; however it has a significant contribution together with distribution parameter uncertainty and physical variability. Also, there is significant interaction between the three sources of uncertainty - physical variability, distribution type and parameter uncertainties.

3.8.2.3 Non-parametric Approach

In addition to the parametric methods, the GP-based non-parametric PDFs are computed for both a and b. These PDFs are then used to compute PDF of y; this

is shown in Fig. 3.24. The non-parametric method does not assume any explicit distribution type or distribution parameters. Thus, the use of a parametric distribution explicitly delineates variability, and epistemic uncertainty (distribution type and parameters) whereas the non-parametric approach represents all the uncertainty in a single distribution.

Similar to the previous example problem, the proposed methodology produces a complete PDF for y, while some of the previously existing solution methodologies result in an interval value for the output y. Kozine and Utkin [94], De Cooman and Troffaes [95], and Ferson and Hajagos [96] produced [0.9, 1.5], [1.0, 1.2], and [0.8, 1.6] respectively. Further, there is considerable agreement between the proposed PDF and the family of PDFs given by Zaman et al. [61].

3.8.3 Discussion of Results

To begin with, it is acknowledged that there is no unique right or wrong answer to problems involving interval uncertainty. Different researchers have pursued different approaches to tackle such problems and this dissertation presents one effective methodology for the analysis of interval uncertainty. It is seen that these different methods have led to comparable solutions, almost similar to one another. Ferson et al. [97] mention four possible reasons for the observed discrepancies among the answers: (1) nesting (due to difference in approaches, one result may be nested in others), (2) differences in truncation, i.e. whether or where the distributions were truncated to finite ranges, (3) numerical approximation error; and (4) different representations of independence.

While the solutions from different methodologies are similar, the proposed methodology has several advantages. It is probabilistic, making it possible to use well-established uncertainty propagation methods such as Monte Carlo simulation, FORM,

SORM, etc. This can provide savings in computational effort, since FORM and SORM typically involve 10 to 20 evaluations of the system response function, and efficient sampling techniques (importance sampling, adaptive sampling, etc.) are available within Monte Carlo Simulation. Second, the proposed methodology is non-parametric, thus making the resulting PDF more loyal to the data than an assumed parametric PDF. Third, different kinds of data can be combined and integrated into a single PDF thereby making the uncertainty representation and propagation simple and straightforward. Fourth, the proposed method provides the entire PDF of the output, which is useful in the context of reliability and risk assessment. Thus, the proposed likelihood-based methodology appears to have strong potential for efficient and effective analysis of interval uncertainty.

3.9 Summary

This chapter proposed statistical methods for the treatment of data uncertainty due to the presence of sparse point and/or interval data to characterize the uncertainty in input quantities. If sufficient point-valued data were available with regard to a particular quantity, then it is possible to precisely identify the distribution type and distribution parameters for the quantity of interest. The presence of sparse point data and interval data leads to uncertainty in both the distribution type and distribution parameters.

The most important concept behind the development of the proposed methods is the construction of the likelihood function for both point and interval data simultaneously. This idea was used for the representation, quantification, and propagation of data uncertainty. This, in turn, led to four research contributions as follows:

1. In Section 3.4, the distribution type of the quantity was assumed to be known

and the uncertainty in the distribution parameters was estimated. This led to a family of distributions, which is computationally expensive for the sake of uncertainty propagation. Hence, the family of distributions was replaced by a single, unconditional distribution which accounts for two types of uncertainty - variability in the quantity (due to assumption of distribution type) and uncertainty in the distribution parameters.

2. Section 3.5 developed a global sensitivity analysis approach to quantify the individual contributions of variability and distribution parameter uncertainty. First, the individual contributions of two types of uncertainty in a single variable was considered. Then, the method was extended to assessing individual contributions of the two types of uncertainty in multiple input variables to the uncertainty in the output of an underlying computational model.

3. Section 3.6 extended the likelihood-based methodology to include uncertainty in distribution type, where multiple competing distribution types were considered. The methods of Bayesian model averaging and Bayesian hypothesis testing were used for this purpose. Further, the sensitivity analysis methodology developed in Section 3.5 was extended to include the uncertainty in the distribution type, and the individual contributions of physical variability and epistemic uncertainty (distribution type uncertainty and distribution parameter uncertainty) were quantified.

4. Section 3.4 and 3.6 both considered parametric distribution types, whereas Section 3.7 extended the likelihood-based methodology to non-parametric distributions. A GP-based non-parametric distribution was developed to directly fit PDFs to point and interval data. Since this approach represents all the types of

uncertainty - variability, distribution type uncertainty and distribution parameter uncertainty - through a single PDF, their individual contributions cannot be calculated.

Future work needs to address other types of epistemic uncertainty such as qualitative information, categorical variables, etc. A probabilistic representation for such quantities may not be straightforward and it is also important to account for such uncertain quantities during uncertainty propagation and integration.

CHAPTER IV

MODEL UNCERTAINTY

4.1 Introduction

Since the late 1960's and early 1970's, significant advancements in the field of computer science and computing technology have encouraged the use of computer models and simulations to solve practical problems in engineering. Solution of a complex partial differential equation to study fluid-structure interaction would have been difficult in the mid 1900's. Today, there are several commercial software packages that can efficiently solve such complicated mathematical equations.

As a result, computational models are increasingly being used to study physical systems in various engineering applications. Additionally, the importance of uncertainty quantification and the impact of the various sources of uncertainty on system response has also been understood. Since the early 2000's, the topic of uncertainty quantification in the system-level performance prediction has gained considerable attention amongst researchers.

Initially, natural variability in the system inputs and parameters was only considered for uncertainty propagation, in order to quantify the uncertainty in the system-level prediction. Due to the growing necessity and desire to analyze and design engineering systems of increasingly complex architectures, it was observed that (1) sufficient data may not be available for uncertainty representation, quantification, and propagation; and more importantly, (2) the ability of the computational model to accurately predict the system performance decreases. While the former issue was

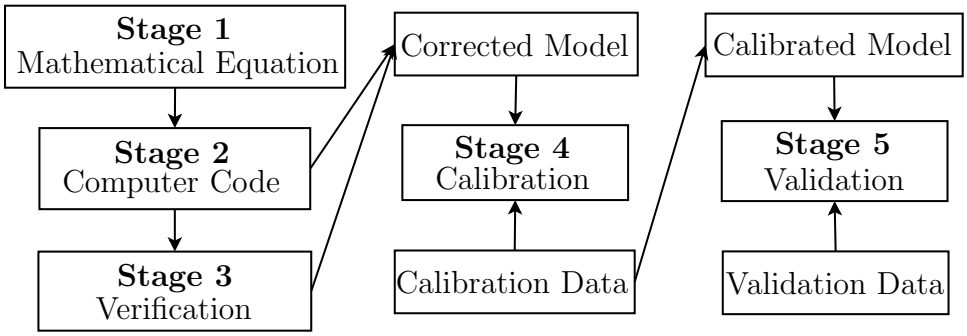


Figure 4.1: Stages of Model Development

addressed in Chapter III, the latter is the result of uncertainty in the model itself. This chapter focuses on the aspect of model uncertainty.

It is essential to quantify the uncertainty in the model, in order to compute the uncertainty in the system-level performance. Helton [1] discusses and illustrates the conceptual and computational basis of Quantifications of Margins and Uncertainties (QMU) in analyses that use computational models to predict the behavior of complex systems. The quantification of model uncertainty is an important component of a QMU analysis that is intimately connected with the assessment, representation, and propagation of uncertainty [1]. First, it is important to quantify the various sources of uncertainty involved in the model development process, and then it is necessary to quantify their combined effect on the system-level prediction.

There are several activities in the development of a model [98], and these activities can be grouped into five steps, as shown in Fig. 4.1. Some of the activities separately delineated by Alvin et al. [98] are collected together in order to facilitate the objectives of this dissertation.

The first step is to develop a conceptual model and construct a mathematical equation (for e.g. a partial differential equation) that represents the model output y as a function of inputs ($\boldsymbol{x}$) and model parameters ($\boldsymbol{\theta}$) as $y = G(\boldsymbol{x}; \boldsymbol{\theta})$. Sometimes, two different physics-based models may be available to represent the same phenomenon.

One model may be more suitable for certain input values, while another may be suitable for other input values. Model selection [99–102] addresses the issue of choosing between two models, in order to compute the output. For example, Section 3.6 discussed a "model selection" problem by considering two competing probability distributions to fit the available data; in the present chapter, the focus is on model uncertainty, and difference is that, now, the models are physics-based models. This dissertation does not address the issue of model selection to study physical systems and processes. It is assumed that a single physics-based model is readily available in mathematical form.

In the second step, a numerical solution procedure is developed to solve the mathematical equation, and this solution procedure is implemented using a computer code. The output of this computer code is the model prediction ($y_c = G_c(\boldsymbol{x}; \boldsymbol{\theta})$); this y_c may be different from y, the true solution of the mathematical equation. When the computer code is used for predictive analysis, it is important to not simply use y_c but to use the actual solution of the mathematical equation.

The third step is the process of model verification, which includes both code verification and solution verification. Code verification focuses on identification of programming errors and debugging computer codes. Solution verification is based on convergence studies and focuses on identification and quantification of solution approximation error, i.e. the difference between y and y_c.

The fourth step is model parameter estimation. The mathematical equation developed in the first step contains some parameters $\boldsymbol{\theta}$ (for example, damping coefficient in a differential equation governing plate deflection under dynamic loading). The model parameters are assumed invariant for different realizations of the input quantities within a range of operating conditions of the system. During model parameter estimation, the values of these parameters need to be estimated based on actual

input-output data. Even if some knowledge regarding the parameters is available, the values of the model parameters need to be adjusted so that the model predictions are in better agreement with experimental data. This adjustment is referred to as model calibration. In the remainder of this dissertation, the terms "model parameter estimation" and "model calibration" are used synonymously.

Having calibrated the model, the fifth step is model validation which answers the question - Is the mathematical equation an accurate representation of reality?. The process of model validation aims to quantify the deviation of the model from reality (referred to as model form error or model discrepancy term or model inadequacy function) and hence, assess the predictive capability of the model.

Hence, there are three major activities related to the quantification of model uncertainty - model verification, calibration, and validation. Note that the steps of verification, validation, and calibration are not necessarily in a fixed sequence; different sequences might be suitable for different problems and there might be iterations between some of the steps. For example, it may be desirable to perform calibration before *and* after validation. The three topics of verification, validation, and calibration are discussed in detail in the following sections.

4.2 Model Verification

The process of verification checks how close the code output is to the true solution of the mathematical equation. As stated earlier in Section 4.1, it is not only sufficient to verify that the two solutions are sufficiently close, but also essential to quantify the solution approximation error, i.e. the difference between the code output and true solution, in order to quantify the uncertainty in the prediction. It is desirable to perform verification before calibration and validation so that the solution approximation errors are accounted for during calibration and validation. Methods for model

verification [103–105] and estimation of solution approximation error [91, 106–109] have been investigated by several researchers.

In general, the solution approximation error is composed of both deterministic and stochastic terms [108]. For example, the discretization error arising in finite element analysis is deterministic, while the surrogate model error that arises as a result of replacing the finite element analysis with a surrogate model is stochastic. In the context of uncertainty propagation, deterministic errors can be addressed by correcting the bias, and the corrected solutions are used to train the surrogate model; the stochastic errors of the surrogate model can be addressed through sampling based on their estimated distributions. As a result, the overall solution approximation error is also stochastic.

The true solution of the mathematical equation can be computed as a function of the model inputs and parameters as $y(\boldsymbol{x};\boldsymbol{\theta}) = y_c(\boldsymbol{x};\boldsymbol{\theta}) + G_{se}(\boldsymbol{x};\boldsymbol{\theta})$. Since $G_{se}(\boldsymbol{x};\boldsymbol{\theta})$ is stochastic, y is stochastic even for given values of $\boldsymbol{x}$ and $\boldsymbol{\theta}$. The remainder of this subsection discusses the estimation of discretization error and surrogate model uncertainty.

4.2.1 Discretization Error

The most common type of solution approximation error is due to discretization in finite element analysis, and methods of convergence analysis [110], a-posteriori error estimation [111], Richardson extrapolation [109, 112, 113], etc. have been studied for the estimation of discretization error. Rebba et al. [91] state that the method of a-posterior error estimation [111, 114, 115] quantifies only a surrogate measure of error to facilitate adaptive mesh refinement, but does not compute the actual discretization error. On the other hand, the method of Richardson extrapolation has been found to come closest to quantifying the actual discretization error [91, 109, 112].

Let h denote the mesh size used in finite element analysis and Ψ the corresponding prediction. Let y denote the "true" solution of the mathematical equation which is obtained as h tends to zero. According to the basic Richardson extrapolation [112], the relation between h and y_c can be expressed as:

$$y = y_c + Ah^p \tag{4.1}$$

In Eq. 4.1, p is the order of convergence, A is the polynomial coefficient. In order to estimate the true solution y, three different mesh sizes ($h_1 < h_2 < h_3$) are considered and the corresponding finite element solutions ($y_c(h_1) = \Psi_1$, $y_c(h_2) = \Psi_2$, $y_c(h_3) = \Psi_3$) are calculated. Eq. 4.1 has three unknowns p, A, and y, which can be estimated based on the three mesh solutions. Mesh doubling/halving is commonly done to simplify the equations. If $r = \frac{h_3}{h_2} = \frac{h_2}{h_1}$, then the discretization error (ϵ_h) and the true solution can be calculated as:

$$\begin{aligned} y &= \Psi_1 - \epsilon_h \\ \Psi_2 - \Psi_1 &= \epsilon_h(r^p - 1) \\ p\ log(r) &= log(\tfrac{\Psi_3 - \Psi_1}{\Psi_2 - \Psi_1}) \end{aligned} \tag{4.2}$$

The solutions Ψ_1 , Ψ_2 , Ψ_3 are dependent on both $\boldsymbol{x}$ and $\boldsymbol{\theta}$ and hence the error estimate ϵ_h and the true solution y are also functions of both $\boldsymbol{x}$ and $\boldsymbol{\theta}$. Since the discretization error is a deterministic quantity, it needs to be corrected for, in the context of uncertainty propagation.

The use of Richardson extrapolation requires uniform meshing and uniform convergence, thereby limiting the applicability of this method in practical finite element analysis. Recently, Rangavajhala et al. [109] has developed a method to overcome these limitations by extending the Richardson extrapolation methodology from a polynomial relation (Eq. 4.1) to a more flexible Gaussian process extrapolation; this

GP is used to extrapolate to $h = 0$ in order to estimate the discretization error. In that case, due to the uncertainty associated with GP interpolation, the discretization error is also stochastic; therefore, the training points (in particular, the output values) for the surrogate model are stochastic, and it is necessary to account for this uncertainty while constructing the surrogate model. Rasmussen [48, 51, 80] discusses constructing GP models when the training point values are stochastic.

4.2.2 Surrogate Model Uncertainty

Another type of solution approximation error arises when the underlying model is replaced with a surrogate model for fast uncertainty propagation and/or model calibration. Surrogate model error is stochastic, even for a given realization of inputs and parameters. As discusses earlier in Section 2.8, different types of surrogate modeling techniques (regression models [27], polynomial chaos expansions [116], radial basis functions [117], support vector machines [46], relevance vector machines [47], Gaussian processes [118]) are available in the literature, and the quantification of the surrogate model error is different for different types of surrogate models. Methods of the quantification of this error (for different surrogate models) are well-established in the literature.

As stated earlier in Section 2.8, this dissertation uses the Gaussian process model as a surrogate to replace expensive computer simulations. There are three important reasons why a Gaussian process model has been used in this research work:

1. The GP model is capable of capturing highly nonlinear relationships that exist between input and output variables without the need for an explicit functional form. Hence, a closed form expression (as in polynomial type regression methods) need not be assumed.

2. For a non-parametric interpolation technique, this method requires fewer sample

points (usually 30 or less) as against methods such as kernel estimation and non-parametric multiplicative regression.

3. A GP model provides a direct estimate of the variance in the output prediction.

Gaussian process interpolation method was explained earlier in Section 2.8 and the Gaussian process prediction (mean and variance) was given by Eq. 2.19. It was emphasized that the choice of training points is important for the construction of the GP model. In Section 2.8, the training points were created based on input-output data of the expensive computer model. Now, the difference is that the training values need to be generated by considering the model parameters ($\boldsymbol{\theta}$) in addition to the inputs ($\boldsymbol{x}$), because the value of the model parameter is also necessary to execute the model and compute the model output.

Once the surrogate model is constructed, the expected value and variance of the Gaussian process prediction can then be used to draw multiple samples for uncertainty analysis, thereby including the effect of surrogate model uncertainty in uncertainty propagation.

4.3 Model Calibration

Model calibration refers to the adjustment of model parameters so that the model output matches well with the field data. The model calibration problem belongs to a wider class of mathematical problems, popularly known as inverse problems. Though this section is titled "model calibration", the focus is on parameter estimation methods, and solutions to inverse problems, as a whole. When a computational model is used to predict the outcome (effect) of a particular phenomenon (cause), this is referred to as the forward problem. The inverse problem makes use of measured

experimental data (effects) to infer the characteristics of the underlying computational model (cause). Alifanov [119] gave a general definition of inverse problems as those that seek to "determine unknown causes based on observation of their effects". Inverse problems are almost synonymously associated with model parameter estimation; in the context of systems with time-dependent output, this problem has also been referred to as system identification [120].

In deterministic analysis, forward problems are usually well-posed, whereas inverse problems are not. According to Hadamard [121], a well-posed problem should have the following properties: (1) a solution exists; (2) the solution is unique; and (2) the solution is stable, i.e. the solution continuously depends on the data, in some reasonable topology.

Since, inverse problems are not well-posed and may have multiple solutions, it becomes essential to assess the confidence associated with the multiple solutions. Further, although most of the formulations of inverse problems directly lead to an optimization problem, it is better to start with a probabilistic formulation, the optimization formulation then appearing as a by-product. Tarantola [122] states that the most simple and generalized theory is obtained when using a probabilistic approach for the solution of inverse problems. Therefore, methods of probability and statistics have been widely used for parameter estimation. Hence, the term "statistical inference" is also used in lieu of parameter estimation.

Several aspects of model calibration and computational methods for parameter estimation are discussed in the following subsections. Some of these methods are well-established in the literature. This dissertation makes two contributions with respect to model calibration. The first contribution addresses numerical issues in the implementation of existing approaches. The second contribution advances the capability of the methods to include different sources of uncertainty and different types

of data situations such as unpaired data and imprecise data. These contributions are interspersed throughout the remainder of this section. Section 4.3.8 is of specific interest, and it discusses several scenarios for model calibration under uncertainty.

4.3.1 The Basic Parameter Estimation Problem

The topic of parameter estimation, i.e. inferring an unobservable (or difficult to measure) quantity through the measurement of a dependent variable, has been a significant topic of interest over several years [12]. A classical example in structural dynamics is the estimation of the damping coefficient based on response measurement.

Consider the computational model $y = G(\boldsymbol{x}; \boldsymbol{\theta})$, where $\boldsymbol{x}$ is the independent input variable and y is the dependent output variable. Point-valued input-output data ($\boldsymbol{x_i}$ vs. y_i; $i = 1$ to n) are assumed to be available. The experimental data is assumed to be unbiased; in other words, $\epsilon_i = y_i - G(\boldsymbol{x_i})$ follows a normal distribution with zero mean. The quantity $\epsilon \sim N(0, \sigma^2)$ is referred to as the fitting error. The goal in parameter estimation is to estimate $\boldsymbol{\theta}$ using the above information.

4.3.2 Least Squares Estimation

The method of least squares is based on minimizing a measure of difference between the model prediction and the observed data. Typically, an error measure $S(\boldsymbol{\theta})$ is computed as:

$$S(\boldsymbol{\theta}) = \sum_{i=1}^{n} (y_i - G(\boldsymbol{x_i}; \boldsymbol{\theta}))^2 \tag{4.3}$$

The so-called least squares estimate of $\boldsymbol{\theta}$ is computed by minimizing the error measure in Eq. 4.3. In order to compute this error measure, note that paired input-output values need to be available. The error measure can be minimized using optimization algorithms, as described by Seber and Wild [71].

This procedure for parameter estimation is also known as non-linear regression. If the model $G(\boldsymbol{x_i};\boldsymbol{\theta})$ is linear with respect to inputs $\boldsymbol{x}$ and parameters $\boldsymbol{\theta}$, then the procedure for estimation of $\boldsymbol{\theta}$ reduces to linear regression and $\boldsymbol{\theta}$ can be calculated analytically, using linear algebra and matrix analysis [27]. Further, Eq. 4.3 assumes that there is no input measurement error and the output measurement error is contained in ϵ. It is also possible to perform regression with measurement errors in both the independent (input) and dependent (output) quantities; this type of regression is commonly referred to as "error-in-variables regression" [123]. Tarantola [122] derives an analytical expression for linear regression with error in variables. In the case of non-linear models, the method of total least squares [124], an extension of Deming regression [125], has been commonly used.

The least squares estimation is a classical statistics-based approach, from a frequentist point of view. The true values of the parameters ($\boldsymbol{\theta}$) are assumed to be deterministic, and the least squares estimate may not coincide with the true value. It can be proved that the least squares estimate tends to the true value, as the data size approaches infinity.

The uncertainty in the least squares estimate is expressed using confidence intervals on the least squares estimate. This confidence interval is calculated at a particular significance level α. Consider the error surface, as shown in Fig. 4.2. (For the sake of graphical illustration, Fig. 4.2 is shown for the case with only one model parameter θ, and not a vector $\boldsymbol{\theta}$).

In Fig. 4.2, the α-level confidence bound is given by the confidence interval $[\theta_{\alpha,min},\theta_{\alpha,max}]$. For a given α, an error value S_α is first defined as:

$$S_\alpha = S(\boldsymbol{\theta^*})(1 + \frac{p}{m-p}F^\alpha_{p,m-p}) \tag{4.4}$$

In Eq. 4.4, F refers to the F-statistic evaluated at significance level α; p refers to the

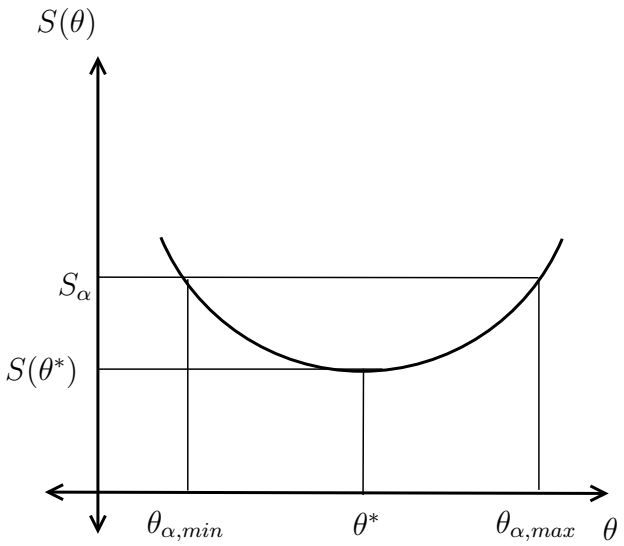


Figure 4.2: Confidence Bounds in Least Squares Analysis

number of parameters being inferred (the length of vector $\boldsymbol{\theta}$, in this case), and n is number of data available of calibration, as defined earlier. The confidence interval of $\boldsymbol{\theta}$ is the region where the condition $S(\boldsymbol{\theta}) \leq S_\alpha$ is satisfied.

In the case of single parameter estimation, an optimization-based procedure is developed in this dissertation to quantify the lower and upper bounds of the confidence interval. Let θ^*, $\theta_{\alpha,min}$ and $\theta_{\alpha,max}$ denote the least squares estimate, lower and upper bounds respectively. This can be accomplished through constraint-based optimization, by maximizing and minimizing θ in order to get the upper and lower bounds respectively, where the constraint is $S(\theta) = S_\alpha$. This is a functionally constrained optimization which may be computationally difficult. Alternatively, the lower and upper bounds of the confidence interval can be computed as follows:

$$\begin{aligned} &\underset{\theta_{\alpha,min}}{\text{Minimize}}\ (S(\theta) - S_\alpha)^2 \text{ s.t. } \theta < \theta^* \\ &\underset{\theta_{\alpha,max}}{\text{Minimize}}\ (S(\theta) - S_\alpha)^2 \text{ s.t. } \theta > \theta^* \end{aligned} \tag{4.5}$$

These optimization problems are functionally unconstrained, one-dimensional and

bounded in one direction. Hence, techniques such as golden section search [126] and parabolic interpolation [127] can be used for quicker solutions. These methods do not use derivatives and hence are computationally efficient as well.

If there are two parameters, then the uncertainty is represented through a two-dimensional confidence region. As the number of parameters increase, it is computationally challenging to quantify the uncertainty associated with the least squares estimate. Further, the above confidence interval/region is not related to the PDF of $\boldsymbol{\theta}$ [9, 71]; since the underlying parameters are assumed to be deterministic, it is meaningless to discuss the PDF of the quantity $\boldsymbol{\theta}$. Hence, it is not possible to propagate the uncertainty in the parameters through another computational model.

4.3.3 The Likelihood Method

The least squares estimation procedure is, fundamentally, an optimization problem. How is it a probabilistic approach, and why is it preferred? It so happens that the least squares estimate maximizes the probability that the given data can actually be observed, under the conditions $\epsilon \sim N(0, \sigma^2)$. This probability, i.e. $P(D|\boldsymbol{\theta})$, where D denotes all the input-output data available, is referred to as the likelihood function of $\boldsymbol{\theta}$, and is denoted as $L(\boldsymbol{\theta})$. The notion of likelihood was formally introduced earlier in Section 2.4.

How to construct the likelihood function for the parameter estimation problem in Section 4.3.1? Assuming that the n pairs of data are independent, the likelihood function can be constructed as:

$$L(\boldsymbol{\theta}) \propto \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{(2\pi)}}\, exp - \left(\frac{(y_i - G(\boldsymbol{x_i}, \boldsymbol{\theta}))^2}{2\sigma^2}\right) \tag{4.6}$$

Note that Eq. 4.6 constructs the likelihood function based on the PDF; the reason

for this was explained earlier in Section 3.4.1. Recall that the likelihood function is meaningful only up to a proportionality constant, and that the likelihood function can be used for inference in the context of both frequentist and subjective probabilities.

4.3.4 Bayesian Inference

The method of Bayesian inference has increasingly gained attention due to the advancements in computing and a substantial increase in computing power. The concept of Bayesian inference was introduced earlier in Section 2.4. The basic idea is to accumulate all prior information in the form of a prior PDF for the parameters $\boldsymbol{\theta}$. This PDF is denoted as $f'(\boldsymbol{\theta})$. The likelihood function $L(\boldsymbol{\theta})$ discussed in Section 4.3.3 is then multiplied with the prior PDF and normalized to calculate the posterior PDF $f''(\boldsymbol{\theta})$.

The standard deviation (σ) of the fitting error (ϵ) can also be estimated simultaneously with the model parameters, by constructing the joint likelihood as:

$$L(\boldsymbol{\theta}, \sigma) \propto \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{(2\pi)}}\ exp - \left[\frac{(y_i - G(\boldsymbol{x_i}, \boldsymbol{\theta}))^2}{2\sigma^2}\right] \tag{4.7}$$

The above equation assumes that measurement errors are not present in the input, since the model is evaluated at the measured input value. Similar to least squares methods, there are also Bayesian approaches for error in variables regression [128]. This is a special case of "inference on mixtures of distributions", a topic widely discussed in the literature [129]; in such a problem, each input-datum and the corresponding output-datum available for calibration are themselves uncertain and hence, need to represented using a probability distribution each. For the sake of simplicity, in the rest of the discussion, measurement errors are assumed to be absent in the independent variables.

The prior for the standard deviation σ can be chosen based on Fisher information criterion [17, 130], as:

$$f'(\sigma) \propto \frac{1}{\sigma} \tag{4.8}$$

Note that this PDF $f'(\sigma)$ is an improper PDF. The joint likelihood is multiplied with the prior and normalized to obtain the joint posterior, denoted by $f''(\boldsymbol{\theta}, \sigma)$. As explained earlier in Section 2.7, MCMC algorithms such as Metropolis sampling [38], Metropolis-Hastings sampling [40], Gibbs sampling [41], and slice sampling [39] are commonly used to generate samples from the joint posterior, without explicitly evaluating the normalizing constant.

The use of MCMC sampling methods requires several hundreds of thousands of evaluations of the model $\boldsymbol{\theta}$. One way to address this challenge is to replace the model $G(\boldsymbol{x}; \boldsymbol{\theta})$ with an inexpensive surrogate. Another way is to explore mathematical methods which provide a significant increase in computational efficiency; in the process, it may be necessary to make a few assumptions which do not alter the result significantly. One such method is described in this chapter in Section 4.3.7; another method is described later in Chapter IX.

4.3.5 Kennedy O'Hagan Framework

The fitting error ϵ is the difference between the model prediction (evaluated at the input measurement) and the corresponding observed output data. The standard deviation of ϵ can be estimated from both frequentist [71] and Bayesian (see Eq. 4.7 in Section 4.3.4) points of view.

The difference between the model prediction and data occurs due two reasons: (1) deviation of model from "reality"; and (2) presence of measurement errors (noise) in the output data. The Kennedy O'Hagan (KOH) framework explicitly delineates these two quantities, by quantifying the so-called model inadequacy function. The

model inadequacy function is also referred to as model discrepancy or model form error. In the KOH approach, the output y is modeled as:

$$y = G(\boldsymbol{x}; \boldsymbol{\theta}) + \delta(\boldsymbol{x}) + \epsilon_m \tag{4.9}$$

where $\epsilon_m \sim N(0, \sigma_m^2)$ is the output measurement error, and $\delta(\boldsymbol{x})$ is modeled as a Gaussian process interpolation. When Bayesian inference is performed, the joint likelihood is constructed for the (1) the model parameters $\boldsymbol{\theta}$; (2) hyper-parameters of model inadequacy function $\delta(\boldsymbol{x})$; and (3) standard deviation (σ_m) of the measurement error (ϵ_m).

If the model $G(\boldsymbol{x}; \boldsymbol{\theta})$ itself needs to be replaced by a Gaussian process surrogate model, then the hyper-parameters of this GP are also jointly estimated along with the above quantities. However, McFarland [52] reports that the uncertainty due to the hyper-parameters of this Gaussian process (the GP that replaces $G(\boldsymbol{x}; \boldsymbol{\theta})$) is negligible compared to the uncertainty in the model parameters, and hence, it may be easier to estimate the hyper-parameters of this GP *before* Bayesian inference.

Based on the discussion in Section 4.1, the quantification model form error is related to model validation. Therefore, the KOH framework combines model calibration and validation activities. The validation method developed in this dissertation, later in Section 4.4 directly calculates the subjective probability that the data supports the model, which is helpful in the context of uncertainty propagation and integration. In this dissertation, the KOH framework is not rigorously implemented; however, such implementation is possible and encouraged.

4.3.6 Regularization

In the beginning of this chapter, it was stated that parameter estimation belongs to a class of inverse problems which are often ill-posed. Sometimes, a regularization technique [8, 12, 131] is used to introduce additional information which then leads to a unique solution; this information may be in terms of penalty for complexity (Occam's razor) or restrictions on the function smoothness (in terms of curvature or second derivative) or bounds on the parameters. The regularization method can be (1) used in a least squares formulation by minimizing the sum of squares of residuals augmented with the above information, i.e. regularization constraint, or (2) applied in a Bayesian framework where the additional information (regularization constraint) is embedded in the prior distribution of the model parameters, which is then multiplied with the likelihood function, and the resultant normalized posterior distribution represents the estimate of the parameters. Regularization procedures are not explicitly implemented in this dissertation; however, such inclusion is encouraged and can be easily facilitated by suitably altering the prior used in Bayesian inference.

4.3.7 Adaptive Integration for Bayesian Inference

This section develops an advanced integration technique to replace the expensive Markov Chain Monte Carlo (MCMC) sampling commonly used to generate samples from the Bayesian posterior, without explicitly evaluating the normalization constant. While one dimensional integrals can be evaluated directly with the proposed technique, multi-dimensional integrals are first transformed to nested one-dimensional integrals and then evaluated through the adaptive quadrature technique.

Consider the case of inferring two variables. Bayesian updating encounters the

evaluation of integrals of the form:

$$k = \int f(x)f(y)L(x,y)dxdy \tag{4.10}$$

Eq. 4.10 consists of a two dimensional integral. However, by the inherent definition of likelihood, this can be converted into two nested one-dimensional integrals. Recalling the definition of likelihood, the likelihood of x and y, $L(x,y)$ is proportional to the probability of observing data conditioned on x and y. Hence, the likelihood of x, i.e. $L(x)$ can be calculated based on the principle of total probability as:

$$L(x) \propto Prob(D|x) \propto \int Prob(D|x,y)f(y)dy \propto \int L(x,y)f(y)dy \tag{4.11}$$

Substituting the expression in Eq. 4.11 into Eq. 4.10,

$$k = \int L(x)f(x)dx = \int \Big(\int L(x,y)f(y)dy \Big) f(x)dx \tag{4.12}$$

Thus, the two-dimensional integral in Eq. 4.10 has been converted into two nested one-dimensional integrals in Eq. 4.12. These one-dimensional integrals can be evaluated using advanced numerical algorithms such as Adaptive Recursive Simpsons Quadrature [132].

Consider any general one-dimensional integral and its approximation using Simpsons rule as:

$$\int_a^b f(x)dx \approx \frac{b-a}{6}\Big(f(a) + 4f\Big(\frac{a+b}{2}\Big) + f(b)\Big) = S(a,b) \tag{4.13}$$

The adaptive recursive quadrature algorithm calls for subdividing the interval of integration (a,b) into two sub-intervals ((a,c) and (c,b), $a \leq c \leq b$) and then, Simpsons

rule is applied to each sub-interval. The error in the estimate of the integral is calculated by comparing the integral values before and after splitting. The criterion for determining when to stop dividing a particular interval depends on the tolerance level ϵ. The tolerance level for stopping may be chosen, for example as [132]:

$$|S(a,c) + S(c,b) - S(a,b)| \leq 15\epsilon \tag{4.14}$$

While the MCMC sampling method may require several hundreds of thousands of evaluations of the model $G(\boldsymbol{x};\boldsymbol{\theta})$, the implementation of the adaptive recursive quadrature algorithm requires about 50 evaluations of each nested integral. Hence, if there are only two parameters of interest, this requires about 2500 evaluations, which is extremely efficient in comparison with the MCMC technique.

It is acknowledged that this method of integration (splitting the integrals and using adaptive recursive Simpsons quadrature) is efficient only when the number of variables is small (≤ 5). If the number of variables is greater than 5, then the MCMC sampling technique is more efficient than the proposed technique.

4.3.8 Model Calibration under Uncertainty

The calibration problem discussion in Sections 4.3.1 - 4.3.5 considered cases when point-valued, paired input-output data are available for calibration. Consider the basic problem with the model $y = G(\boldsymbol{x};\boldsymbol{\theta})$; this basic problem can be expanded to include different features, as explained below. These different features include situations where different sources of uncertainty may be present in the model and/or data. It is expected that the uncertainty in the model parameters increase with the presence of additional sources of uncertainty. The goal is to quantify this uncertainty in the model parameters. From hereon, the concept of likelihood and the Bayesian

approach are pursued rigorously, since the least squares method can neither rigorously account for the various sources of uncertainty nor calculate the PDF of the model parameters.

4.3.8.1 Additional Sources of Uncertainty

The model prediction may also depend on some other quantities $\boldsymbol{\alpha}$ which are known to be uncertain and cannot be measured while collecting calibration data. Further, such quantities $\boldsymbol{\alpha}$ are not calibrated because it may not be physically meaningful to calibrate them.

Hence, the model is represented as $y = G(\boldsymbol{x}; \boldsymbol{\theta}, \boldsymbol{\alpha})$, and the uncertainty in $\boldsymbol{\alpha}$ is denoted in terms of its PDF $f(\boldsymbol{\alpha})$. Similar to Section 4.3.1, point-valued input-output data ($\boldsymbol{x_i}$ vs. y_i; $i = 1$ to n) are assumed to be available to calibrate $\boldsymbol{\theta}$; however, now the difference is that the model prediction at input $\boldsymbol{x_i}$ is uncertain. Further, the PDF of the model prediction at $\boldsymbol{x_i}$ is not statistically independent of that at $\boldsymbol{x_j}$ ($i \neq j$) because the same PDF $f(\boldsymbol{\alpha})$ is used for both cases, even though the measurements ($\boldsymbol{x_i}$ vs. y_i) are independent of each other.

The likelihood function of $\boldsymbol{\theta}$ can be constructed to include the uncertainty in $\boldsymbol{\alpha}$, as:

$$L(\boldsymbol{\theta}) = \int L(\boldsymbol{\theta}, \boldsymbol{\alpha}) f(\boldsymbol{\alpha}) d\boldsymbol{\alpha} \tag{4.15}$$

The likelihood $L(\boldsymbol{\theta})$ is used in Bayesian inference to compute the posterior PDF of $\boldsymbol{\theta}$. In Eq. 4.15, the likelihood function $L(\boldsymbol{\theta}, \boldsymbol{\alpha})$ is calculated as:

$$L(\boldsymbol{\theta}, \boldsymbol{\alpha}) \propto \prod_{i=1}^{n} \left[\frac{1}{\sigma\sqrt{(2\pi)}}\, exp - \left(\frac{(y_i - G(\boldsymbol{x_i}, \boldsymbol{\theta}, \boldsymbol{\alpha}))^2}{2\sigma^2} \right) \right] \tag{4.16}$$

In Eq. 4.16, the likelihood is calculated only for a particular value of $\boldsymbol{\alpha}$ and hence

the independence between the measurements can be used to multiply individual likelihoods.

Note that in Eq. 4.15, the calculation of likelihood is a multi-dimensional integration, where the number of dimensions is equal to the number of uncertain quantities in $\boldsymbol{\alpha}$. When this likelihood is substituted in Bayesian, the calculation of posterior involves multi-dimensional integration, where the number of dimensions is equal to the number of calibration parameters in $\boldsymbol{\theta}$. Hence, this requires a nested multi-dimensional integration. This issue of the presence of additional sources of uncertainty is also discussed in detail, later in Chapter IX.

4.3.8.2 Interval Data for Calibration

Consider the calibration problem with the model $y = G(\boldsymbol{x}; \boldsymbol{\theta})$. Sometimes, the data for calibration is available in the form of intervals. For the sake of illustration, consider m intervals, $[\boldsymbol{a_i}, \boldsymbol{b_i}]$ at the input level and corresponding $[c_i, d_i]$ at the output level. How to construct the likelihood for this case? Censored data, often available in reliability analysis [79], is a special case of interval data. Suppose that the number of cycles to failure is measured in reliability testing; if the specimen does not fail until N cycles, then the number of cycles to failure is a censored interval, i.e. (N, ∞).

The likelihood-based approach for representation of interval data (developed earlier in Section 3.4) cannot be applied here because, if all the intervals were represented using a combined PDF, then the "orderedness" or "correspondence" between the input and output pairs would be lost. Hence, each interval has to be treated separately. Each interval is represented using a uniform distribution on the interval $[\boldsymbol{a_i}, \boldsymbol{b_i}]$ and the corresponding PDF is denoted as $f(\boldsymbol{\chi_i})$ ($i = 1$ to m; $\boldsymbol{a_i} \leq \boldsymbol{\chi_i} \leq \boldsymbol{b_i}$). Note that $\boldsymbol{a_i}$, $\boldsymbol{b_i}$, and $\boldsymbol{\chi_i}$ are vectors; each member of this vector corresponds to a member in the input vector $\boldsymbol{x_i}$.

These PDFs can be used to construct the likelihood function for $\boldsymbol{\theta}$, in terms of the individual likelihoods, as:

$$L(\boldsymbol{\theta}) \propto \prod_{i=1}^{m} L_i(\boldsymbol{\theta}) \tag{4.17}$$

where the individual likelihood $L_i(\boldsymbol{\theta})$ can be calculated by including the PDF $f(\boldsymbol{\chi_i})$ as:

$$L_i(\boldsymbol{\theta}) = \int L_i(\boldsymbol{\chi_i}, \boldsymbol{\theta}) f(\boldsymbol{\chi_i}) d\boldsymbol{\chi_i} \tag{4.18}$$

The likelihood $L_i(\boldsymbol{\chi_i}, \boldsymbol{\theta})$ in Eq. 4.18 is calculated for one realization of the input $\boldsymbol{\chi_i}$, as:

$$L_i(\boldsymbol{\chi_i}, \boldsymbol{\theta}) \propto \int_{y=c_j}^{y=d_j} \left(\frac{1}{\sigma\sqrt{(2\pi)}} \; exp - \left(\frac{(y - G(\boldsymbol{\chi_i}, \boldsymbol{\theta})^2}{2\sigma^2} \right) \right) dy \tag{4.19}$$

Note that Eq. 4.19 uses a CDF to account for the interval data as against the PDF in Eq. 4.6. This aspect is similar to the treatment of interval data in Section 3.4.1.

4.3.8.3 Partially Characterized Data for Calibration

Consider the calibration problem with the model $y = G(\boldsymbol{x}; \boldsymbol{\theta})$. Typically, in an experiment, the value of the independent variable (input) is selected, the experiment is performed, and the corresponding measurement of the dependent variable (output) is used for calibration; such measurements are well-characterized. Sometimes, it may not be possible to conduct experiments in such a way that the input and the output measurements have one-to-one correspondence. In other words, the input measurements are conducted independent of the output measurements; such measurements are referred to be "partially characterized" or "uncharacterized" in this dissertation. Further, each of the measurements (input and/or output) may be point-valued or an interval. How to construct the likelihood for this case?

Consider m point data x_i ($i = 1$ to m) and n intervals $[a_i, b_i]$ ($i = 1$ to n), available for a particular input x; note the vector of inputs is not considered here. Since

there is no one-to-one correspondence between the input and output measurements, all of the input measurements can be aggregated.

From a frequentist point of view, one possible approach is to construct a composite PDF, as:

$$f_X(x) = \frac{1}{m+n}(\sum_{i=1}^{m} \delta(x - x_i) + \sum_{i=1}^{n} U_X(a_i, b_i)) \tag{4.20}$$

In Eq. 4.20, $\delta(.)$ refers to the Dirac delta function, and $U_X(a_i, b_i)$ refers to the PDF of a uniform distribution defined on the interval $[a_i,b_i]$, as shown in Eq. 4.21.

$$U_X(a_i, b_i) = \begin{cases} \frac{1}{b_i - a_i} & \text{if } a_i \leq x \leq b_i \\ 0 & \text{else} \end{cases} \tag{4.21}$$

Thus, each point data is represented as a Dirac delta function, and each interval is represented using a uniform distribution. The input PDF $f_X(x)$ is expressed as a weighted sum of all these distributions, where each weight is equal to $\frac{1}{m+n}$, assuming that each data (point or interval) is weighed equally.

Alternatively, from a subjectivist point of view, the methods for data uncertainty quantification developed in Chapter III can be used to construct the PDF $f_X(x)$ for the input x; the parametric methods in Sections 3.4 and 3.6 or the non-parametric method in Section 3.7 can be used for this purpose.

The above procedure for the calculation of $f_X(x)$ is repeated for all the input variables which are uncharacterized, and the joint PDF of the inputs is denoted as $f_{\boldsymbol{X}}(\boldsymbol{x})$.

This PDF can be used in uncertainty propagation to compute the model prediction as a function of the parameter $\boldsymbol{\theta}$, using uncertainty propagation methods discussed in Section 2.5. Let $f_Y(y|\boldsymbol{\theta})$ denote the corresponding model prediction; note that this is

computed as a function of $\boldsymbol{\theta}$, in order to facilitate the construction of the likelihood function $L(\boldsymbol{\theta})$. This likelihood is constructed using the output data available.

At the output level, consider p point data y_i ($i = 1$ to p) and q intervals $[c_i,\ d_i]$ ($i = 1$ to q). Similar to the previous sections, the likelihood is calculated using the PDF value for point data and CDF values for interval data as:

$$L(\boldsymbol{\theta}) \propto \left[\prod_{i=1}^{p}\Big(\int f(z=y_i|y) f_Y(y|\boldsymbol{\theta})dy\Big)\right]\times \left[\prod_{j=1}^{q}\Big(\int\Big(\int_{z=c_i}^{z=d_i} f(z|y)dz\Big) f_Y(y|\boldsymbol{\theta})dy\Big)\right] \tag{4.22}$$

In Eq. 4.22, z is simply used as a dummy variable, and $f(z|y)$ is calculated similar to Eq. 4.6, as:

$$f(z|y) = \frac{1}{\sigma\sqrt{(2\pi)}}\ exp - \left(\frac{(z-y)^2}{2\sigma^2}\right) \tag{4.23}$$

As in the previous sections, the likelihood function can be used in Bayesian inference in order to compute the posterior PDF of $\boldsymbol{\theta}$.

4.3.8.4 Calibration under Uncertainty: Synopsis

Conventionally, model calibration has considered paired input-output measurements for calibration. In this dissertation, several scenarios for model calibration are considered:

1. Additional sources of uncertainty
2. Interval data for calibration
3. Uncharacterized input-output data

The methods proposed to address the above situations can be easily extended to address situations where unpaired data, interval data, and other sources of uncertainty

are all simultaneously present. A numerical example is presented in Section 4.3.10 to illustrate the proposed methods. Further, the topic of model calibration under uncertainty will again be revisited in Chapter IX.

4.3.9 Estimating $\boldsymbol{\theta}$ versus Distribution Parameters of $\boldsymbol{\theta}$

Sections 4.3.1 - 4.3.8 considered the estimation of the model parameter $\boldsymbol{\theta}$ based on input-output data. Sometimes, it may be known that the model parameters are naturally varying quantities, which follow a particular distribution type. In such cases, it may be necessary to estimate the distribution parameters of the model parameters $\boldsymbol{\theta}$. Let the PDF of $\boldsymbol{\theta}$ be denoted by $f_{\boldsymbol{\theta}}(\boldsymbol{\theta}|\boldsymbol{P_\theta})$. How to construct the likelihood function if the aim is to estimate the distribution parameters ($\boldsymbol{P_\theta}$) of $\boldsymbol{\theta}$ rather than $\boldsymbol{\theta}$ itself?

The likelihood must be a function of those parameters which need to be estimated. Hence, in this case, it is desired to calculate the likelihood function as $L(\boldsymbol{P_\theta})$. The methods developed in Sections 4.3.1 - 4.3.8 can be used to calculate the likelihood of the model parameter, i.e. $L(\boldsymbol{\theta})$ This function can be calculated using the principle of conditional probability, as:

$$L(\boldsymbol{P_\theta}) \propto \int L(\boldsymbol{\theta}) f_{\boldsymbol{\theta}}(\boldsymbol{\theta}|\boldsymbol{P_\theta}) d\boldsymbol{\theta} \tag{4.24}$$

Then the likelihood function $L(\boldsymbol{P_\theta})$ can be used in Bayes theorem (Eq. 2.7) to estimate the entire PDFs of the distribution parameters ($\boldsymbol{P_\theta}$). This leads to a family of distributions for the model parameter $\boldsymbol{\theta}$. If desired, the principle of total probability can be used to compute a single PDF for the model parameter $\boldsymbol{\theta}$, similar to that in Eq. 3.8.

4.3.10 Application: Energy Dissipation in a Lap Joint

Sections 4.3.1 - 4.3.9 discussed several aspects of model calibration and parameter estimation. Now, an engineering application problem in the area of mechanical engineering is chosen for illustrating the proposed methods. For the sake completion, calibration is performed using both classical statistics-based methods and likelihood-based Bayesian methods.

4.3.10.1 Description of the Problem

This example deals with the calibration of the Smallwood model [91, 133, 134], which is used to predict the energy dissipation due to friction at a lap joint in a mechanical component. This model predicts the dissipation energy (D_E) per cycle at the joint when the component is subjected to an impact harmonic force of amplitude F.

The hysteresis curve (force vs. displacement graph) for the lap joint comprises of two symmetrical portions. The energy loss in the joint under one cycle of sinusoidal loading is found by integrating the area under the hysteresis curve and analytically derived as:

$$D_E = k_n(\frac{m-1}{m+1})z^{m+1} \tag{4.25}$$

In Eq. 4.25, k_n is a non-linear stiffness term, m is the exponent term, and z is the displacement amplitude which is obtained by solving:

$$2F = kz - k_n z^m \tag{4.26}$$

In Eq. 4.26, k refers to a linear stiffness term. The objective is to calibrate the non-linear stiffness parameter (k_n) using the available input-output data. Data is available on the inputs - force (F), linear stiffness (k), and non-linear exponent (m) - and the

output - dissipated energy (D_E). For a given force, the dissipated energy can be measured, and hence there is correspondence and ordered pairing between the input force (F) and output energy (D_E). There are five such measurements as shown in Table 4.1.

Table 4.1: Calibration Data: Force vs. Dissipated Energy

Force (F) (in lbf)	Dissipated Energy (Z) (in $lbf \times in$)
60	5.30×10^{-5}
120	2.85×10^{-4}
180	7.78×10^{-4}
240	1.55×10^{-3}
320	2.50×10^{-3}

Let F_j ($j = 1$ to 5) and Z_j ($j = 1$ to 5) denote the five force values and five energy values in Table 4.1. A different symbol Z has been used for the output measurement in order to avoid confusion with the symbol D_E used for model prediction. The variables Z and D_E are related through the fitting error ($\epsilon \sim N(0, \sigma^2)$), as:

$$f(Z|D_E) = \frac{1}{\sigma\sqrt{(2\pi)}}\ exp - \left(\frac{(Z - D_E)^2}{2\sigma^2}\right) \tag{4.27}$$

Two other inputs, linear stiffness (k) and non-linear exponent (m) are not measured in correspondence with the force measurement and hence are not paired with the output measurement as well. For the sake of illustrating the methods, it is assumed that this unpaired data has come from other independent sources (other experiments, subject matter experts, etc.)

The data on linear stiffness (k) is available as: three intervals ([1160000, 1180000], [1155000, 1170000], [1160000, 1170000]) and one point value (1173000). All measurements are in lbf/in.

The non-linear exponent (m) is known to have a normal distribution with mean = 1.23 (no units) and coefficient of variation = 0.06. Thus, this numerical example

features several types of uncertainty - (1) additional source of uncertainty (m); (2) interval data (k); and (3) uncharacterized data (k).

For the purpose of model calibration, Eq. 4.25 and Eq. 4.26 can be expressed together as

$$D_E = G(F, k, m; k_n) \tag{4.28}$$

Eq. 4.28 is obtained by eliminating z from Eq. 4.25 and Eq. 4.26. The various steps in the calibration procedure are:

1. Represent each unpaired input using a PDF. There are two unpaired inputs. The linear stiffness (k) is given by three intervals and one point value, and the nonlinear exponent (m) is known to be normal. Thus, there is a well-defined PDF $f_m(m)$ in the latter case whereas the PDF $f_k(k)$ needs to be constructed for the former case.

2. Once $f_m(m)$ and $f_k(k)$ are known, then the model in Eq. 4.28 needs to be evaluated for each value of the paired input, i.e. for each of the force values in Table 4.1.

3. In the least squares approach, an error measure is computed as a function of the calibration parameter (k_n) and minimized whereas in the likelihood approach, the likelihood function of the calibration parameter (k_n) is computed and maximized.

4. The uncertainty in least squares estimation is expressed through confidence intervals, and the likelihood function is used in Bayesian inference to calculate the entire PDF of the calibration parameter (k_n).

These four steps are explained in detail below.

4.3.10.2 Least Squares Approach

In the first step, $f_m(m)$ is known to be a normal density function with mean = 1.36 and coefficient of variation = 0.05.

The PDF $f_k(k)$ is constructed as a composite PDF - a weighted sum of three uniform PDFs ([1160000, 1180000], [1155000, 1170000], [1160000, 1170000]) and one Dirac delta PDF centered at the available point data (1173000). Uniform weights are used for each of these 4 data; hence each of these data is assigned a probability of 0.25, and cumulative probabilities of 0.25, 0.50, 0.75, and 1 respectively. The algorithm for generating samples from this composite PDF is:

1. Generate random number. Call it "temp".

2. If temp < 0.25, Generate a sample for k from the uniform distribution bounded on [1160000, 1180000]. End. To generate next sample, Go to Step 1.

3. Else if temp < 0.50, Generate a sample for k from the uniform distribution bounded on [1155000, 1170000]. End. To generate next sample, Go to Step 1.

4. Else if temp < 0.75, Generate a sample for k from the uniform distribution bounded on [1160000, 1170000]. End. To generate next sample, Go to Step 1.

5. Else, the sample for k is 1173000. End. To generate next sample, Go to Step 1.

Using the above generated samples, the PDF of k is generated and shown in Fig. 4.3.

This completes the first step.

In the second step, the PDF of the model prediction is estimated as a function of F_j ($j = 1$ to 5) and k_n. It needs to be a function of F because the input measurement (F) and the output measurement (Z) have a one-one correspondence, as in Table 4.1. It needs to be a function of k_n because this quantity is the calibration parameter

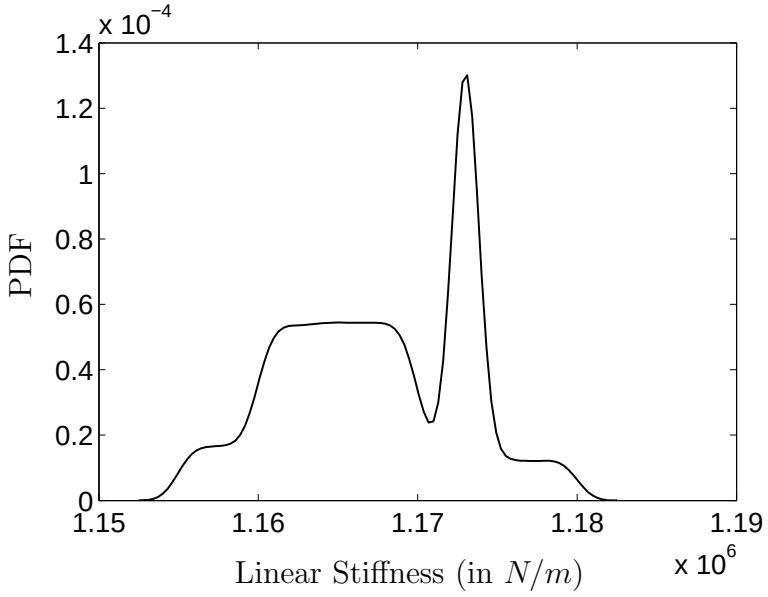


Figure 4.3: Composite PDF of Linear Stiffness

and the squared error measure should be a function of k_n. Uncertainty propagation methods (Section 2.5) are then used to compute the PDF of model prediction for each ordered input measurement; this PDF is denoted as $f_{D_{E_j}}(D_{E_j}|F_j, k_n)$. In the third step, this PDF $f_{D_{E_j}}(D_{E_j}|F_j, k_n)$ is compared (so as to compute an error measure $S_j(k_n)$) with the corresponding output measurement data denoted by Z_j, which is a point value.

The comparison of a PDF with a point value is not straightforward; this PDF $f_{D_{E_j}}(D_{E_j}|F_j, k_n)$ cannot be effectively used in the least squares approach, and this is one reason why the least squares approach is not used in the remainder of this dissertation. Nevertheless, for the sake of completeness, this example problem is completed as follows.

The error metric $S(k_n)$ is computed by measuring the difference between the measurement Z_j and the expectation of the model prediction, $i.e. E(D_{E_j}|F_j, k_n)$, as:

$$S(k_n) = \sum_{j=1}^{5} S_j(k_n) = \sum_{j=1}^{5} \Big(E(D_{E_j}|F_j, k_n) - Z_j \Big)^2 \tag{4.29}$$

(If the output measurement were an interval instead of a point value, then a uniform PDF can be considered on this interval and this uniform PDF is compared with $f_{D_{E_j}}(D_{E_j}|F_j, k_n)$ to compute $S_j(k_n)$. A distance metric such as the area metric [135], Kullback-Leibler Divergence [136], Hellinger distance [137], Bhattacharyya distance [138] etc. can be used to compute the "distance" or "difference" between two PDFs. It must be noted that the choice of a uniform distribution is an additional assumption. In the likelihood/Bayesian approach, these difficulties are easily overcome using the inherent definition of likelihood.)

This error metric $S(k_n)$ is minimized and the calibration parameter (k_n) is estimated. Further, the uncertainty in this estimate is calculated using the F-statistic, as explained earlier in Section 4.3.2. The least squares estimate is found to be 6.62 x $10^5 lbf/in$ and the 95% confidence interval is calculated to be [6.59 x $10^5 lbf/in$, 6.66 x $10^5 lbf/in$].

4.3.10.3 Likelihood Approach

In the likelihood approach, the first step and the third step are different from the least squares approach; the second step of computing the PDF $f_{D_{E_j}}(D_{E_j}|F_j, k_n)$ using uncertainty propagation is the same.

In the first step, the PDF $f_k(k)$ is computed using the non-parametric likelihood-based technique described earlier in Section 3.7, and the resulting PDF is shown in Fig. 4.4.

This PDF is estimated by solving the optimization in Eq. 3.30; the domain of [1155000, 1180000] is discretized into 10 equal parts, and the density values at these discretization points are estimated so that the likelihood for the given data (3 intervals and one point) is maximum. Note that this is a non-parametric PDF constructed using the Gaussian process interpolation method based on the PDF values at the

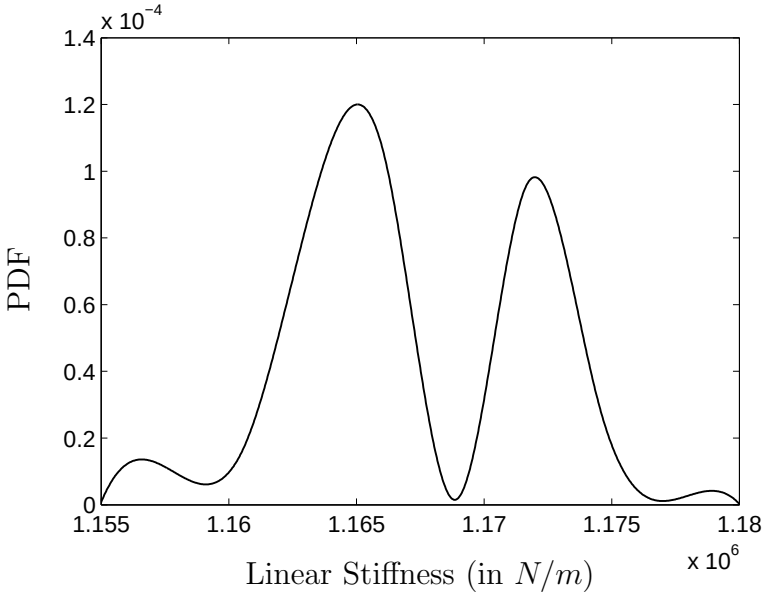


Figure 4.4: Likelihood-based PDF of Linear Stiffness (k)

10 discretization points. Also, this PDF does not make an additional assumption (uniform distribution within each interval which was assumed in Fig. 4.3). This completes the first step in model calibration.

As mentioned earlier, the second step is the same as in the least squares approach, and the PDFs $f_{D_{E_j}}(D_{E_j}|F_j, k_n)$ ($j = 1$ to 5) are calculated using uncertainty propagation.

In the third step the likelihood function is calculated using the methods developed in Section 4.3.8. Then, this likelihood function can be maximized; further the uncertainty in this estimate can be calculated using Bayes theorem as explained earlier in Eq. 2.7. A non-informative uniform distribution for the prior of k_n, and the resulting posterior PDF is shown in Fig 4.5.

The maximum likelihood estimate is 660690 lbf/in; the 95% probability bounds are given by [617160 lbf/in, 689060 lbf/in].

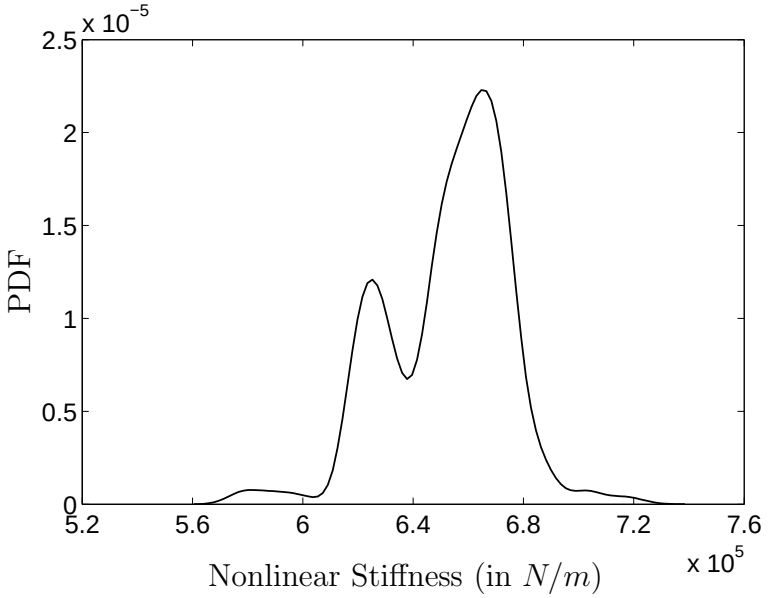


Figure 4.5: Estimated PDF of Nonlinear Stiffness (k_n)

4.3.10.4 Discussion

To begin with, it is acknowledged that there is no single correct answer to the model calibration problem with imprecise and unpaired input-output descriptions. First, there are several procedures in the literature to treat interval data and each procedure may lead to a different answer. Ferson et al. [97] discuss in detail several aspects of interval data treatment. The least squares procedure is able to address interval data only by assuming a uniform distribution within each interval. As stated earlier, this assumption is questionable. The likelihood approach does not make any such assumption and hence, is faithful to the data, as far as possible. Second, the issue of correspondence between input and output measurements directly affects the uncertainty in the calibrated quantity. Note that the multi-modal behavior of the input k is reflected in the PDF of the calibrated k_n in Fig. 4.5.

Also, the estimates of uncertainty obtained from the least-squares based approach and the likelihood approach need to be interpreted differently. The likelihood approach alone can provide the entire PDF of the calibrated k_n and the least squares

approach cannot; it only provides confidence intervals which need to be interpreted differently from probability bounds [9]. A 95% confidence interval for a model parameter is interpreted as follows: "Suppose that the data collection process was repeated 100 times. Hence, 100 different least squares estimates, and 100 corresponding confidence intervals can be calculated. Then, 95 out of 100 intervals will contain the *true* estimate of the model parameter".

4.3.11 Model Calibration: Summary

Model calibration and parameter estimation are important issues in the context of model uncertainty quantification. Model parameter uncertainty significantly affects the predictive capability of the model and in many situations, the model output may be highly sensitive to the uncertainty in the model parameters. Conventional model calibration requires well-characterized, point-valued, paired input-output data for model parameter estimation. In this dissertation, the focus was on model calibration under uncertainty, especially in the presence of unpaired and imprecise data. Though least squares-based methods and Bayesian methods were discussed, the emphasis was on the use of the latter methods. The proposed methods were illustrated using non-linear structural dynamics experimental data on energy dissipation due to friction in a lap joint, under impact loading.

The topic of model calibration will be revisited in Chapter V and Chapter IX. In Chapter V, these methods are used in crack growth analysis in order to estimate the crack growth parameters, and thereby predict the crack growth as a function of number of cycles. In Chapter IX, model calibration in the presence of multiple sources of uncertainty will be considered in detail; a rigorous computational methodology will be developed to apportion the model parameter uncertainty to the multiple sources

of uncertainty. The sensitivity of model parameters to the calibration data also will be studied.

4.4 Model Validation

Model validation refers to the process of quantifying the extent to which the computational model under study is supported by available experimental data. The model validation procedure measures the extent of agreement between the model prediction and the experimental data. A visual comparison, usually referred to as "graphical validation", though valuable, is inadequate in many cases [104, 139]. Such an approach is only qualitative and cannot explicitly account for the different sources of uncertainty.

Oberkampf and Barone [140] explain the need for rigorous quantitative validation metrics which can be perceived as computable measures that can compare model predictions and experimental results over a range of input (or control) variables to sharpen the assessment of computational accuracy. An important aspect of model validation is the rigorous, explicit treatment of multiple sources and different types of uncertainty. The various types of uncertainty can be broadly classified into two types - aleatory and epistemic. In the context of validation, both the model inputs and the experimental evidence are uncertain. A rigorous approach to model validation should explicitly account for the various sources of uncertainty such as physical variability, information uncertainty, measurement error, solution approximation errors (for example, discretization error) etc. and develop a robust metric that can quantitatively judge the performance of the model and assess the confidence in the model prediction.

Model validation under aleatory uncertainty has been studied by several researchers and there are methods available in the literature [91, 92, 105, 140–145] to solve this

problem. These methods are based on the use of statistical techniques such as confidence intervals [140], normalizing difference between experimental observation and model prediction (normalized difference [105, 141], area metric [135]), classical statistics-based hypothesis testing [142], Bayesian hypothesis testing [91, 92, 144, 145], and model reliability analysis [143]. These methods dealt with validation under aleatory uncertainty in the model inputs and parameters, and included measurement errors and model solution approximation errors, i.e. the result of verification.

While the method of classical hypothesis testing is simply based on comparison of the mean (and variance) of model prediction vs. the mean (and variance) of the experimental data, the method of Bayesian hypothesis testing not only allows the comparison of entire distributions of model prediction and experimental data, but also provides a systematic approach to account for various types of uncertainty. The method of Bayesian hypothesis testing has been pursued by Mahadevan and co-workers [91, 92, 143, 145–147] at Vanderbilt University, and it can directly quantify the extent to which the computational model is supported by the experimental data. This method has been applied to the validation of both reliability prediction models [92, 145] and performance prediction models [91, 146, 147]. In this approach, the validation metric is the Bayes factor, which is the ratio of the likelihood that the model is correct to the likelihood that the model is incorrect. Jiang and Mahadevan [147] showed how the threshold Bayes factor for model acceptance can be derived based on a risk vs. cost trade-off, thus facilitating physically meaningful decision making. Further, using Bayes theorem, the Bayes factor can be directly used to compute the probability the model is correct, relative to the alternate hypothesis.

Though it is clear that the Bayesian hypothesis testing approach is suitable for the purpose of model validation, this method cannot be used directly when the model

inputs and validation data are quantities with epistemic uncertainty. The inclusion of epistemic uncertainty (sparse and interval data) was difficult because while probabilistic methods have been commonly used for model validation, interval data has been primarily addressed using non-probabilistic methods such as evidence theory [64], convex models of uncertainty [67], Zadeh's extension principle [69], fuzzy set theory [70], etc. Hence, some researchers have resorted to non-probabilistic methods for model validation in the presence of epistemic uncertainty [148, 149]. These methods are based on interval analysis, and hence computationally expensive, especially when both aleatory and epistemic uncertainty are present; in this case, Roy and Oberkampf [149] recommend nesting probabilistic analysis (to address aleatory uncertainty) within interval analysis (to address epistemic uncertainty) for uncertainty propagation and model validation. This approach significantly increases the computational effort.

In this dissertation, two probabilistic methodologies are pursued for the purpose of model validation under uncertainty. The first approach is based on the concept of Bayesian hypothesis testing, and the second approach is based on the concept of model reliability. In the past, these two methods have only considered aleatory uncertainty. The contribution of this dissertation is to extend these two approaches to include the different sources of epistemic uncertainty (sparse point data and interval data), and address several scenarios such as unpaired input-output data, time series data, etc. for model validation.

Consider the computational model $y = G(\boldsymbol{x}; \boldsymbol{\theta})$, where $\boldsymbol{x}$ refers to the inputs and $\boldsymbol{\theta}$ refers to the model parameters, and y is the model output. Output measurement data, available for validation, is denoted by z. Note that z and y refer to the same physical quantity (for example, displacement, temperature, etc.). Different symbols have been used to differentiate the actual available data (z) from the corresponding

model output (y). The aim of model validation is to quantify the extent of agreement between z and y.

During the process of model validation, it is also necessary to account for the output measurement error (ϵ_o). Conventionally, this term is assumed to be normally distributed as N(0, σ_o^2). Given the model prediction, the probability of observing the data z can be calculated as being proportional to:

$$f(z|y) = \frac{1}{\sqrt{2\pi}\sigma_o} exp\bigg(- \frac{(y-z)^2}{2\sigma_o^2} \bigg) \tag{4.30}$$

During validation, it is assumed that the model parameters ($\boldsymbol{\theta}$) are known, either deterministically or in terms of the PDF ($f_{\boldsymbol{\theta}}(\boldsymbol{\theta})$). The Bayesian hypothesis testing and model reliability approaches are both investigated for model validation, by different scenarios (interval data, uncharacterized data, etc.).

4.4.1 Bayesian Hypothesis Testing

This section introduces the Bayesian hypothesis testing procedure for model validation, and extends this methodology to include epistemic uncertainty and account for both well-characterized and uncharacterized data. Additionally, time-series data for validation is also considered.

The fundamental premise of Bayesian hypothesis testing is based on two contradicting hypotheses: a null hypothesis (H_0) and an alternate hypothesis (H_1). The goal is to quantify the extent of support from experimental data (D) to each of these hypotheses.

In Section 3.6.2, the Bayesian hypothesis method was used to quantify the support from data to multiple competing distribution types (denoted by M_1 and M_2 in Section 3.6.2). In the context of model validation, the null hypothesis (H_0) refers to the event that the model is correct (i.e. the data supports the model), and the alternate

hypothesis (H_1) refers to the event that the model is not correct (i.e. the data does not support the model). These two hypotheses can be mathematically written as:

$$\begin{aligned} H_0 &\rightarrow y = z \\ H_0 &\rightarrow y \neq z \end{aligned} \tag{4.31}$$

If prior probabilities ($P(H_0)$ and $P(H_1)$) are known for each of these hypotheses, then Bayes theorem can be used to update their probabilities based on the available data (D), as [91]:

$$\frac{P(H_0|D)}{P(H_1|D)} = \frac{P(D|H_0)}{P(D|H_0)} \frac{P(H_0)}{P(H_1)} \tag{4.32}$$

In Eq. 4.32, $P(H_0|D)$ and $P(H_1|D)$ refer to the posterior probabilities of the events H_0 and H_1. In the context of model validation, $P(H_0|D)$ is the posterior probability that the model is correct.

The first term on the right hand side of Eq. 4.32 is referred to as the Bayes factor, denoted by B [17].

$$B = \frac{P(D|H_0)}{P(D|H_0)} \tag{4.33}$$

The Bayes factor is the ratio of the likelihood of the null hypothesis to that of the alternative hypothesis. If the prior probabilities of the null and alternate hypotheses are equal, then the posterior probability that the model is correct, i.e. $P(H_0|D)$ can be expressed in terms of the Bayes factor as:

$$P(H_0|D) = \frac{B}{B+1} \tag{4.34}$$

If the Bayes factor is greater than one, then the data D favors the event "model is correct". Higher the Bayes factor, higher is the likelihood of the null hypothesis, and better is the agreement between the model prediction and the experimental data.

4.4.1.1 Calculation of Bayes Factor

In order to implement the Bayesian hypothesis testing procedure, the key is the calculation of the Bayes factor. This calculation depends on how the data is available for validation. Consider the conventional case, where n point-valued input-data ($\boldsymbol{x_i}$ vs. y_i; $i = 1$ to n) is available for validation. Assume that the n data were collected through independent experiments. If the model parameters ($\boldsymbol{\theta}$) are deterministic, then $P(D|H_0)$ is based on the model prediction, as:

$$P(D|H_0) \propto \prod_{i=1}^{n} f(z = y_i|G(\boldsymbol{x_i};\boldsymbol{\theta})) \tag{4.35}$$

If the model parameters ($\boldsymbol{\theta}$) are uncertain, and described using the PDF $f_{\boldsymbol{\theta}}(\boldsymbol{\theta})$, then this uncertainty must be included in the calculation of $P(D|H_0)$, as:

$$P(D|H_0) \propto \int \Big(\prod_{i=1}^{n} f(z = y_i|G(\boldsymbol{x_i};\boldsymbol{\theta})) \Big) f_{\boldsymbol{\theta}}(\boldsymbol{\theta}) d\boldsymbol{\theta} \tag{4.36}$$

In Eq. 4.36, note that the product over n data points is computed for a single realization of $\boldsymbol{\theta}$, and then, the product is integrated over the space of $\boldsymbol{\theta}$. Hence, the integration accounts for the parameter uncertainty, propagates this uncertainty to the output, and the product is nested within the integral. The computation of Eq. 4.36 requires a multi-dimensional integration. The method of Monte Carlo integration [150] can be used to evaluate this integral. It may be tempting to compute $P(D|H_0)$ as:

$$P(D|H_0) \propto \prod_{i=1}^{n} \left(\int f(z = y_i|G(\boldsymbol{x_i};\boldsymbol{\theta})) f_{\boldsymbol{\theta}}(\boldsymbol{\theta}) d\boldsymbol{\theta} \right) \tag{4.37}$$

Eq. 4.37 computes $P(D|H_0)$ by nesting the integral within the product. This implies that uncertainty propagation is performed before multiplication. Therefore, the

inherent assumption is that the model prediction PDFs $f_Y(\boldsymbol{x_i})$ and $f_Y(\boldsymbol{x_j})$ are independent of each other. This is equivalent to the statement that the model prediction at a particular input value is independent of the model prediction at another input value. This is obviously *wrong* since the same model parameters ($\boldsymbol{\theta}$) and the PDF $f_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ are used to compute $f_Y(\boldsymbol{x_i})$ and $f_Y(\boldsymbol{x_j})$. This is not related to the fact that the experiments and hence, output measurements are statistically independent of each other. Hence, it is mathematically incorrect to use Eq. 4.37 to compute $P(D|H_0)$. On the other hand, in Eq. 4.36, the product is computed only for one realization of $\boldsymbol{\theta}$, and therefore, the multiplication is justified through the assumption of independent experiments.

The next important question is regarding the computation of $P(D|H_1)$. In order to compute this probability, it is necessary to assume an alternate hypothesis for the prediction quantity. Since the hypothesis H_1 means that the model is not correct, all subjective knowledge (regarding the prediction quantity) that can be gained without using the model is used to construct the alternate hypothesis. Hence, a PDF for y is assigned under the alternate hypothesis. This PDF is denoted as $f_Y^A(y|\boldsymbol{x}, H_1)$; note that this PDF is conditioned on the input value since the output y is conditioned on the input value, as $y(\boldsymbol{x})$. In practice, this "alternate" PDF can be chosen to be uniform, and the bounds can be obtained from subject matter experts. Then, $P(D|H_1)$ can be calculated as:

$$P(D|H_1) \propto \prod_{i=1}^{n} \Big(\int f(z = y_i|y) f_Y^A(y|\boldsymbol{x_i}, H_1) dy \Big) \tag{4.38}$$

One disadvantage in the use of the Bayesian hypothesis method is the requirement of this alternate PDF. Another important issue is whether the PDFs $f_Y^A(y|\boldsymbol{x_i}, H_1)$ and $f_Y^A(y|\boldsymbol{x_j}, H_1)$ are statistically independent of each other. Since the alternate PDF needs to be assumed, it is practically difficult to assume an alternate PDF for each

$\boldsymbol{x_i}$ and it is further difficult to assume PDFs with statistical dependence. In fact, Eq. 4.38 is valid only if the alternate PDF at one input value is independent of the alternate PDF at another input value. Alternatively, if statistically dependent PDFs can be assumed, then this dependence must be accounted for, as:

$$P(D|H_1) \propto \int \Big(\prod_{i=1}^{n} f(z = y_i|y) \Big) f_Y^A \Big(y(\boldsymbol{x_1}) \ ... \ y(\boldsymbol{x_n}) | H_1 \Big) dy(\boldsymbol{x_1}) \ ... \ dy(\boldsymbol{x_n}) \quad (4.39)$$

The numerical examples discussed in this dissertation do not assume such dependent PDFs and hence, Eq. 4.39 is neither implemented nor considered in the remainder of this dissertation.

Then, the Bayes factor can be computed using Eq. 4.33 and the posterior probability that the model is corrected can be calculated using Eq. 4.34. The following discussions present model validation situations when point-valued paired input-output data is not available. The difference lies in the calculation of $P(D|H_0)$ and $P(D|H_1)$.

4.4.1.2 Interval Data for Validation

Sometimes, interval data are available for validation. For example, consider m intervals, $[\boldsymbol{a_i}, \boldsymbol{b_i}]$ at the input level and corresponding $[c_i, d_i]$ at the output level. The likelihood-based methodology for the representation of interval data (developed earlier in Section 3.4) cannot be applied here because, if all the intervals were represented using a combined PDF, then the "correspondence" between the input and output pairs would be lost. Hence, each interval has to be treated separately. Each interval is represented using a uniform distribution on the interval $[\boldsymbol{a_i}, \boldsymbol{b_i}]$ and the corresponding PDF is denoted as $f(\boldsymbol{\chi_i})$ ($i = 1$ to m; $\boldsymbol{a_i} \leq \boldsymbol{\chi_i} \leq \boldsymbol{b_i}$). Note that $\boldsymbol{a_i}$, $\boldsymbol{b_i}$, and $\boldsymbol{\chi_i}$ are vectors; each member of this vector corresponds to a member in the input vector $\boldsymbol{x_i}$.

Consider the case where the model parameters ($\boldsymbol{\theta}$) are uncertain and the PDF is

given by $f_{\boldsymbol{\theta}}(\boldsymbol{\theta})$. Then, the probability $P(D|H_0)$ can be computed as:

$$P(D|H_0) \propto \int \left(\prod_{i=1}^{n} \left(\int \left(\int_{z=c_i}^{z=d_i} f(z|G(\boldsymbol{\chi_i};\boldsymbol{\theta}))dz \right) f(\boldsymbol{\chi_i}) d(\boldsymbol{\chi_i}) \right) \right) f_{\boldsymbol{\theta}}(\boldsymbol{\theta}) d\boldsymbol{\theta} \quad (4.40)$$

Since the input measurements are intervals, an alternate PDF of prediction quantity needs to be available for each available input interval, as $f_Y^A(y|[\boldsymbol{a_i}, \boldsymbol{b_i}], H_1)$. Then, the probability $P(D|H_1)$ can be computed as:

$$P(D|H_1) \propto \prod_{i=1}^{n} \left(\int \left(\int_{z=c_i}^{z=d_i} f(z|y)dz \right) f_Y^A(y|[\boldsymbol{a_i}, \boldsymbol{b_i}], H_1)dy \right) \quad (4.41)$$

Then, the Bayes factor can be computed using Eq. 4.33 and the posterior probability that the model is corrected can be calculated using Eq. 4.34.

4.4.1.3 Partially Characterized Data

Sometimes, it may not be possible to conduct experiments in such a way that the input and the output measurements have one-to-one correspondence. In other words, the input measurements are conducted independent of the output measurements; such measurements are referred to be "partially characterized" or "uncharacterized" in this dissertation, similar to Section 4.3.8.3. Further, each of the measurements (input and/or output) could be point-valued or an interval. How to compute the Bayes factor in this case?

Consider m point data x_i ($i = 1$ to m) and n intervals $[a_i, b_i]$ ($i = 1$ to n), available for an input x. Since there is no one-to-one correspondence between the input and output measurements, all of the input measurements can be aggregated. The classical statistics-based method developed in Section 4.3.8.3 may be used for this purpose; the basic concept is to compute the composite PDF by treating each point data as Dirac delta PDF and each interval as a uniform PDF. Alternately, the the

methods for data uncertainty quantification developed in Chapter III can be used to construct the PDF $f_X(x)$ for the input x; the parametric methods in Sections 3.4 and 3.6 or the non-parametric method in Section 3.7 can be used for this purpose.

The above procedure for the calculation of $f_X(x)$ is repeated for all the input variables which are uncharacterized, and the joint PDF of the inputs is denoted as $f_{\boldsymbol{X}}(\boldsymbol{x})$. The model prediction is then computed using uncertainty propagation; the PDFs $f_{\boldsymbol{X}}(\boldsymbol{x})$ and $f_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ are propagated through the model to compute the PDF of the model output $f_Y(y|H_0)$; note that this PDF is conditioned only on the model and not on inputs or parameters because their uncertainty is included in this PDF.

At the output level, consider p point data y_i ($i = 1$ to p) and q intervals $[c_i,\ d_i]$ ($i = 1$ to q). Then $P(D|H_0)$ is calculated as:

$$\begin{aligned} P(D|H_0) \propto & \left[\prod_{i=1}^{p} \Big(\int f(z = y_i|y) f_Y(y|H_0) dy \Big) \right] \times \\ & \left[\prod_{j=1}^{q} \Big(\int \Big(\int_{z=c_i}^{z=d_i} f(z|y) dz \Big) f_Y(y|H_0) dy \Big) \right] \end{aligned} \tag{4.42}$$

In the case of unpaired data, only one PDF $f_Y^A(y|H_1)$ needs to be assumed for the alternate hypothesis. Then, $P(D|H_0)$ is calculated as:

$$\begin{aligned} P(D|H_1) \propto & \left[\prod_{i=1}^{p} \Big(\int f(z = y_i|y) f_Y^A(y|H_1) dy \Big) \right] \times \\ & \left[\prod_{j=1}^{q} \Big(\int \Big(\int_{z=c_i}^{z=d_i} f(z|y) dz \Big) f_Y^A(y|H_1) dy \Big) \right] \end{aligned} \tag{4.43}$$

Then, the Bayes factor can be computed using Eq. 4.33 and the posterior probability that the model is corrected can be calculated using Eq. 4.34.

4.4.1.4 Time Series Data

Consider a time-dependent model $y(t) = G(\boldsymbol{\theta})$, and n data are available for validation of this model, as $[t_i,\ y_i]$ ($i = 1$ to n). Without loss of generality, it can be assumed that the data is increasingly arranged with time, i.e. $t_i < t_{i+1}$.

If the data came from n different physical components, then they do *not* belong to the same time series. In that case, validation can be performed similar to Section 4.4.1.1, by simply treating t as another input to the model G; in other words, $y(t) = G(\boldsymbol{\theta})$ is written as $y = G(t, \boldsymbol{\theta})$.

If all the n data points came from the same time series, then the equations in Section 4.4.1.1 are not applicable for validation. Consider the computation of $P(D|H_0)$, where $D = D_1 \cap D_2\ .\ .\ .\ D_n$, and D_i represents $[t_i,\ y_i]$. The equations in Section 4.4.1.1 would imply that $y(t_i)$ is independent of $y(t_j)$, for a given realization of $\boldsymbol{\theta}$. This is not true in the case of time series data. Therefore, two types of dependences need to be accounted for:

1. $y(t_i)$ is dependent on $y(t_j)$ (even for one realization of $\boldsymbol{\theta}$) due to the presence of a time series.

2. The PDFs $f_Y(y(t_i))$ and $f_Y(y(t_j))$ are statistically dependent because the model parameters $\boldsymbol{\theta}$ (and the corresponding $f_{\boldsymbol{\theta}}(\boldsymbol{\theta})$) are used for prediction in both the cases.

In order to rigorously account for these dependencies, a multi-step procedure for validation is proposed. The first step is to calculate the Bayes factor for D_1, as a function of $\boldsymbol{\theta}$, as:

$$B_1(\boldsymbol{\theta}) \propto \frac{f(y_1|y(t_1))f_Y(y(t_1)|\boldsymbol{\theta})}{\int f(y_1|y(t_1))f_Y^A(y(t_1)dy(t_1)} \tag{4.44}$$

From hereon, the Bayes factor is computed successively for the next data D_2, by

conditioning on the previous data D_1. In general, for $i > 1$,

$$B_i(\boldsymbol{\theta}) \propto \frac{f(y_i|y(t_i))f_Y(y(t_i)|\boldsymbol{\theta}, y(t_{i-1}) = y_{i-1})}{\int f(y_i|y(t_i))f_Y^A(y(t_i)|y(t_{i-1}) = y_{i-1})dy(t_i)} \tag{4.45}$$

Hence, a Bayes factor is calculated for every D_i by conditioning on D_{i-1}. As a result, the overall Bayes factor is a product of all the Bayes factors, as:

$$B(\boldsymbol{\theta}) = \prod_{i=n}^{n} B_i(\boldsymbol{\theta}) \tag{4.46}$$

Note that this Bayes factor is still conditioned on the model parameters $\boldsymbol{\theta}$. The principle of conditional probability is then used to calculate the unconditional Bayes factor, as:

$$B = \int B(\boldsymbol{\theta}) f_{\boldsymbol{\theta}}(\boldsymbol{\theta}) d\boldsymbol{\theta} \tag{4.47}$$

Then the Bayes factor can be used to compute the posterior probability that the model is correct as in Eq. 4.34.

4.4.1.5 Bayesian Hypothesis Testing: Summary

In this dissertation, the method of Bayesian hypothesis testing has been extended to include different situations for model validation under uncertainty, including (1) parameter uncertainty; (2) interval data; (3) partially characterized data; and (4) time series data. One disadvantage of this method is that the alternate hypothesis requires the choice of a PDF for the prediction quantity; this choice may be subjective. It must be noted that the Bayes factor metric significantly depends on the choice of the alternate PDF, and hence this metric is a relative measure of the support from data to the model vs. alternate hypothesis. If the alternate hypothesis adequately represents the event "the model G is not correct", then the posterior probability

calculated using the Bayes factor is an accurate measure of the probability that the model is correct.

4.4.2 Reliability-based Metric

This section discusses the model reliability approach for model validation, where the model is considered to reliable if difference between the model prediction and the experimental data is less than a tolerance level. Though Rebba and Mahadevan [143] originally presented this approach, the model validation metric was computed based only on the mean of the model prediction and the output (experimental) data. In this dissertation, a computational methodology is developed to compute the model reliability metric based on the entire probability distributions of the model prediction and the output experimental data. Different scenarios such as well-characterized data, uncharacterized data, interval data, and time series data are considered.

The basic concept is to accept the model if the difference between the model prediction (y) and the observed data (z) is less than a tolerance limit (δ). Let M denote the event of accepting the model; then

$$M \rightarrow |z - y| \leq \delta \equiv z - \delta \leq y \leq z + \delta \tag{4.48}$$

Then $P(M)$ quantifies the probability that the model is correct. This metric is an absolute metric, in comparison with the Bayes factor metric. Since Eq. 4.48 is similar to reliability analysis, this metric $P(M)$ is referred to as the model reliability metric. The method is developed for both well-characterized and uncharacterized data.

4.4.2.1 Well-characterized Data

Consider the computational model $y = G(\boldsymbol{x}, \boldsymbol{\theta})$. Consider the conventional case for model validation, where n point-valued input-data ($\boldsymbol{x_i}$ vs. y_i; $i = 1$ to n) is available for validation. Assume that the n data were collected through independent experiments.

First consider a particular input-output pair ($\boldsymbol{x_i}$ vs. y_i). Based on Eq. 4.48, the probability that the model is correct for this data ($P(M_i)$), is calculated as:

$$P(M_i) = \int \left(\int_{z=y_i-\delta}^{z=y_i+\delta} f(z|G(\boldsymbol{x_i}, \boldsymbol{\theta}))dz \right) f_{\boldsymbol{\theta}}(\boldsymbol{\theta})d\boldsymbol{\theta} \tag{4.49}$$

In Eq. 4.49, the concept is to propagate the PDF $f_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ through the model and then use the entire PDF of the model prediction in Eq. 4.48 to calculate the model reliability.

Once these probabilities are computed for each data point, it may be tempting to multiply them because the data are independently collected. This is not correct, because the *same* PDF $f_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ is propagated through the model at multiple input values ($\boldsymbol{x_i}$; $i = 1$ to n), and the corresponding model predictions are not independent. Therefore, it is wrong to multiply all $P(M_i)$'s. In order to account for this dependence, the model reliability metric, $P(M)$ is calculated as:

$$P(M) = \int \left(\prod_{i=1}^{n} \left(\int_{z=y_i-\delta}^{z=y_i+\delta} f(z|G(\boldsymbol{x_i}, \boldsymbol{\theta}))dz \right) \right) f_{\boldsymbol{\theta}}(\boldsymbol{\theta})d\boldsymbol{\theta} \tag{4.50}$$

In Eq. 4.50, the product is computed for a given realization of $\boldsymbol{\theta}$, and justified because of the assumption that the data are collected through independent experiments.

Alternatively, in Eq. 4.49, the model reliability can be computed conditioned on $\boldsymbol{\theta}$, as $P(M_i|\boldsymbol{\theta})$. Then, these conditional probabilities for each data point can be multiplied to compute the overall conditional probability ($P(M|\boldsymbol{\theta})$). Finally, the

overall conditional probability can be calculated as:

$$P(M) = \int P(M|\boldsymbol{\theta}) f_{\boldsymbol{\theta}}(\boldsymbol{\theta}) = \int \Big(\prod_{i=1}^{n} P(M_i|\boldsymbol{\theta}) \Big) f_{\boldsymbol{\theta}}(\boldsymbol{\theta}) \boldsymbol{\theta} \tag{4.51}$$

Therefore, the key is the computation of the conditional probability $P(M_i|\boldsymbol{\theta})$ for each input-output pair.

For a particular data (j^{th} data), if the input measurement is an interval ($\boldsymbol{a_j}$, $\boldsymbol{b_j}$), then the interval is approximated with a uniform distribution $f(\boldsymbol{\chi_j})$ ($\boldsymbol{a_j} \leq \boldsymbol{\chi_j} \leq \boldsymbol{b_j}$). Note that $\boldsymbol{a_j}$, $\boldsymbol{b_j}$, and $\boldsymbol{\chi_j}$ are vectors; each member of this vector corresponds to a member in the input vector $\boldsymbol{x_j}$. Then the conditional model reliability for this data is calculated as:

$$P(M_j|\boldsymbol{\theta}) = \int_{\boldsymbol{\chi_j}=\boldsymbol{a_j}}^{\boldsymbol{\chi_j}=\boldsymbol{b_j}} \left(\int_{z=y_j-\delta}^{z=y_j+\delta} f(z|G(\boldsymbol{\chi_j}, \boldsymbol{\theta}))dz \right) f(\boldsymbol{\chi_j}) d(\boldsymbol{\chi_j}) \tag{4.52}$$

Suppose that for a particular data (k^{th} data), the output measurement is an interval (c_k, d_k). The next important question is - "what does this mean in the context of model reliability in Eq. 4.48 ?" If the model prediction is y and the output measurement is an interval (z_1, z_2), the the model is acceptable if $z_1 - \delta \leq y \leq z_2 + \delta$. Therefore, the conditional model reliability for this data is calculated as:

$$P(M_i|\boldsymbol{\theta}) = \int_{z=c_k-\delta}^{z=d_k+\delta} f(z|G(\boldsymbol{x_i}, \boldsymbol{\theta}))dz \tag{4.53}$$

While Eq. 4.52 considers interval input measurement and point-valued output measurement, Eq. 4.53 considers point-valued input measurement and interval output measurement. It is straightforward to extend these two equations to the case when both the input and output measurements are intervals.

4.4.2.2 Partially Characterized Data

The previous discussion considered paired input output measurement. Similar to Sections 4.3.8.3 and 4.4.1.3, partially characterized input-output data is now considered.

Consider m point data x_i ($i = 1$ to m) and n intervals $[a_i, b_i]$ ($i = 1$ to n), available for an input x. Then a single PDF $f_X(x)$ is used to represent all of this information; this PDF can be constructed using the classical statistics-based approach in Section 4.3.8.3 or likelihood-based approach in Chapter III.

The above procedure for the calculation of $f_X(x)$ is repeated for all the input variables which are uncharacterized, and the joint PDF of the inputs is denoted as $f_{\boldsymbol{X}}(\boldsymbol{x})$. The model prediction is then computed using uncertainty propagation; the PDFs $f_{\boldsymbol{X}}(\boldsymbol{x})$ and $f_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ are propagated through the model to compute the PDF of the model output $f_Y(y)$; note that this PDF is conditioned only on the model and not on inputs or parameters because their uncertainty is included in this PDF.

Similarly, the output measurements are available in the form of p point data y_i ($i = 1$ to p) and q intervals $[c_i, d_i]$ ($i = 1$ to q). A similar procedure is used to represent this information using a PDF; let $f_Z(z)$ denote this PDF. Then, $P(M)$ can be easily calculated based on Eq. 4.48. A simple Monte Carlo analysis can be used for this purpose, since this calculation does not require any further evaluations of the model G.

4.4.2.3 Reliability-based Metric: Summary

The model reliability approach addresses model validation from the perspective of reliability analysis, where the model is considered to be acceptable if the difference between the model prediction and the experimental observation is less than a tolerance

limit. This approach is not relative, in comparison with the Bayesian hypothesis testing method.

There is one disadvantage with the model reliability approach for model validation. The product in Eq. 4.50 may tend to zero as the number of validation points increases. This is because, as the number of points increases, it is increasingly difficult to satisfy the condition in Eq. 4.48 for *all* points. Therefore, a large tolerance level δ may need to be chosen in order to achieve a significantly high model reliability metric.

4.4.3 Application: Energy Dissipation in a Lap Joint

This section illustrates the proposed methods for validating the Smallwood model [133, 134, 143, 146], which is used to predict the energy dissipation due to friction at a lap joint in a mechanical component. The dissipated energy is dependent on the (1) applied force; (2) linear stiffness k; (3) exponent m; and (4) non-linear stiffness k_n.

The Smallwood model was earlier calibrated (k_n being the parameter estimated) in Section 4.3.10, by assuming interval data on k, and a probability distribution for m. Now, the focus is on validation given the uncertainty in the model parameters. So, the proposed validation methods are illustrated for given PDFs of the parameters, as given in Table 4.2. Model validation under imprecise, unpaired data is considered in another numerical example, in Section 4.4.4.

Table 4.2: Mean and Standard Deviation of Model Parameters

Parameter	m	$log_{10}(k_n)$	k
Mean	1.23	5.61	1172655
SD	0.06	0.28	13700

In Table 4.2, m has no unit, k and k_n have unit of lbf/in. The correlation matrix of these three parameters is given in Table 4.3.

Table 4.3: Correlations Between Model Parameters

m	$log_{10}(k_n)$	k
1	0.96	0.62
0.96	1	0.41
0.62	0.41	1

Note that three parameters are correlated, and hence $f_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ in Section 4.4.2 must account for this correlation.

There are five ordered input-output (force-energy) pairs as shown in Table. 4.4. The Bayesian hypothesis testing procedure is used to calculate the Bayes factor and the posterior probability that the model is correct for each ordered pair. It must be noted that the value of the Bayes factor and the Bayesian posterior probability depend on the choice of the alternate hypothesis $f_y(y|H_1)$. A uniform distribution, corresponding to 95% probability bounds of model prediction was chosen, for the purpose of illustration. Similarly, the model reliability metric is computed for each ordered input-output pair, by considering a tolerance level equal to 5% of the observed experimental value.

Table 4.4: Smallwood Model: Validation Data and Results

Force F (in lbf)	Energy D (in $lbf \times in$)	Mean of Prediction	Bayesian Posterior $P(H_0\|D)$	Model Reliability $P(M)$
60	5.30×10^{-5}	5.70×10^{-5}	0.57	0.42
120	2.85×10^{-4}	2.70×10^{-4}	0.60	0.60
180	7.78×10^{-4}	6.75×10^{-4}	0.33	0.35
240	1.55×10^{-3}	1.30×10^{-3}	0.30	0.21
320	2.50×10^{-3}	2.50×10^{-3}	0.62	0.74

From Table 4.4, it can be seen that whenever the data is closer to the prediction, the value of the validation metric is higher. Overall, the model is only marginally acceptable in three cases based on the Bayesian metric (where $P(H_0|D) > 0.5$), and in 2 cases based on model reliability metric (where $P(M) > 0.5$). However, none of the evidence is able to rule completely in favor of the Smallwood model.

4.4.4 Application: Heat Conduction

This section presents another numerical example that illustrates the proposed methodologies for model validation in the presence of uncharacterized data. For the purpose of illustration, different types of uncertainty - sparse point data, interval data and distribution parameter uncertainty - are assumed to exist simultaneously in the model inputs. Further, the experimental observations may be point data or interval data.

Consider the steady state heat transfer in a thin wire of length L, with thermal conductivity k, and convective heat coefficient β. The temperature at midpoint of the wire needs to be predicted. For the sake of illustration, it is assumed that this problem is essentially one dimensional and that the solution can be obtained from the following boundary value problem [91].

$$\begin{aligned} -k\frac{\partial^2 T}{\partial T^2} + \beta T &= Q(x) \\ T(0) &= T_0 \\ T(L) &= T_L \end{aligned} \tag{4.54}$$

Assume that the heat source $Q(x) = 25(2x - L)^2$. Rebba et al. [91] assumed that the temperatures at the ends of the wire are both zero ($T_0 = T_L = 0$). This is an ideal scenario and this example considers uncertainty in the boundary condition, i.e. the temperatures at the ends of the wire ($T_0 = T_L = 0$) are assumed to be normally distributed with statistics N(0, 0.1). This is an example of aleatory uncertainty or physical variability (in the boundary condition).

Suppose that the probability distribution of the conductivity of the wire k is described by an expert as normal but with uncertain distribution parameters. For the sake of illustration, it is assumed that the mean follows a normal distribution

with statistics $N\ (5, 0.2)$ and the standard deviation follows a normal distribution with mean equal to 1 and standard deviation = 0.1. These values are in SI units, i.e. $Wm^{-1}/^{\circ}C$.

Suppose that a distribution is not available for the convective heat coefficient β. Instead it is described using two intervals, [0.5, 0.60] and [0.45, 0.48] and two additional point data are available: 0.58 and 0.52. These values are in SI units, i.e. $Wm^{-2}/^{\circ}C$.

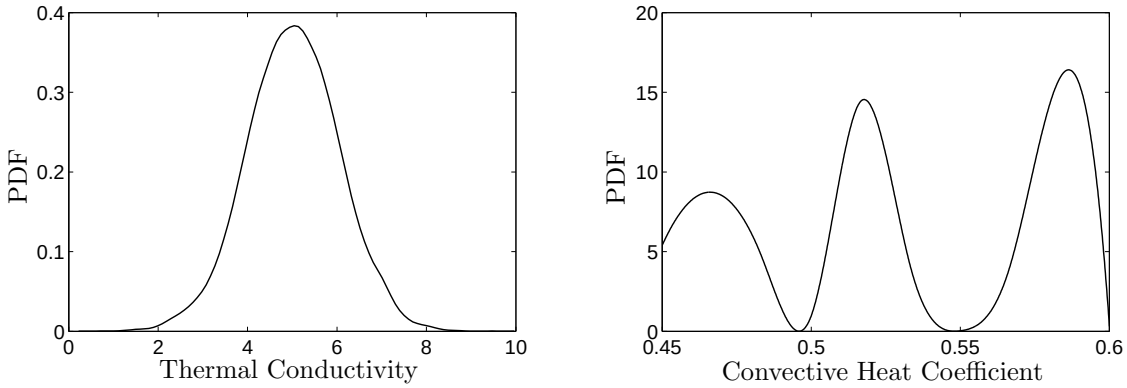
The length of the wire is assumed to be deterministic, L = 4 m. Let the purpose of the model be to predict the temperature at the middle of the wire, i.e. at x = 2.0, in Celcius.

For the purpose of illustration, assume that output measurement data consists of two points and two intervals, i.e. 15, 16, [16 18], and [15.5 17]. It is required to assess whether this experimental evidence supports the numerical model in Eq. 4.54. The output experimental error is assumed to have zero mean, and standard deviation equal to 5% of the observed data.

The first step is to construct the PDFs of conductivity of the wire k, and convective heat coefficient β. The former has distribution parameter uncertainty and hence the PDF is conditioned on the parameters. The unconditional PDF can be estimated using Eq. 3.8. The latter is described using sparse point and interval data, and the PDF is calculated using the non-parametric approach developed in Section 3.7. These two PDFs are shown in Fig. 4.6a and Fig. 4.6b respectively.

Then, the PDF of the model prediction is computed using uncertainty propagation, by propagating all the parameter uncertainties, and physical variability in the boundary conditions, and the resulting PDF is shown in Fig. 4.7.

Then the non-parametric PDF of the observed data (based on two point data and two interval data) is also computed using the non-parametric approach developed in



(a) PDF of k (b) PDF of β

Figure 4.6: PDFs of Model Parameters

Section 3.7. The resultant PDF is shown in Fig. 4.8. Note that this PDF is used only in the model reliability approach and not in the Bayesian hypothesis testing approach. In the latter, the point data and interval data are directly included while constructing the likelihood function.

First, the model is validated using the Bayesian hypothesis testing method. A uniform distribution bounded on the interval [12, 22] is chosen for the PDF of y under the alternative hypothesis, i.e. $f_y(y|H_1)$. The Bayes factor, and the Bayesian posterior probability ($P(H_0|D)$) are computed to be 4 and 0.80 respectively. Then, the method is validated using the model reliability approach; for the sake of illustration, three different tolerance levels are considered, as shown in Table 4.5.

Table 4.5: Heat Conduction Problem: Model Validation Results

Tolerance δ (in Celcius)	Model Reliability $P(H_0\|D)$
1	0.3
2	0.58
3	0.71

Though both the methods - Bayesian hypothesis testing and model reliability approach - are used to calculate "the probability that the model is correct", it is

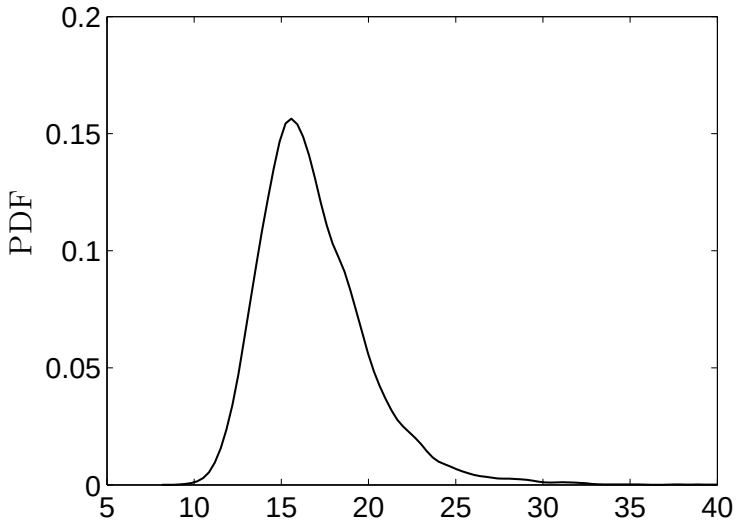


Figure 4.7: PDF of the Model Prediction

important not to compare them, because they are different from each other. In Bayesian hypothesis testing, the choice of $f_y(y|H_1)$ is subjective, and affects the model validation result; in fact, the Bayesian posterior probability is simply a relative measure of support from data to the model vs. alternate hypothesis. Therefore, it is necessary to be sufficiently informed about this PDF $f_y(y|H_1)$, in order to obtain accurate results. On the other hand, there are no assumptions in the model reliability metric, and hence it is likely to yield a more objective metric for model validation.

4.4.5 Model Validation: Summary

Model validation is important in the context of uncertainty quantification, to measure the extent of agreement between the model prediction and experimental data. In this dissertation, two approaches - Bayesian hypothesis testing and reliability-based method - were investigated for model validation under uncertainty. Both of these methods were extended to include several types of uncertainty and different types of data scenarios such as interval data, uncharacterized data, and time series data.

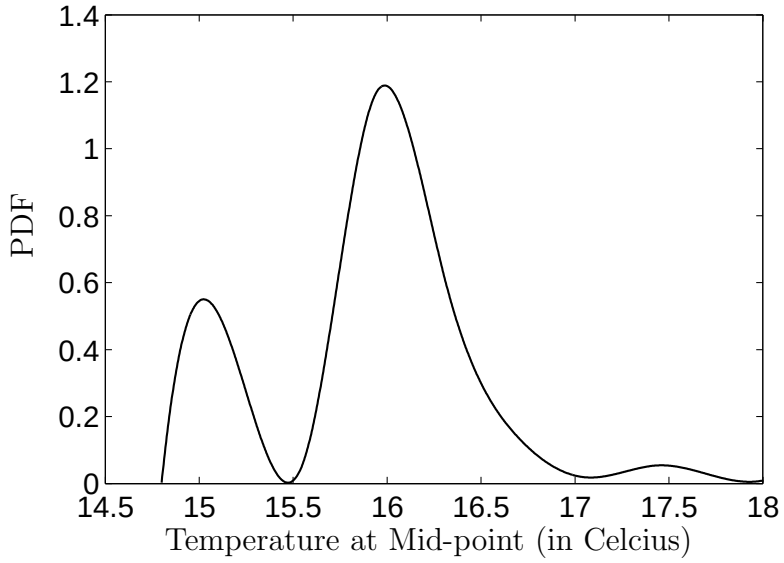


Figure 4.8: PDF of the Observed Data

The Bayesian hypothesis testing computes a Bayes factor metric for model validation which is relative to the choice of the alternate hypothesis. On the other hand, the model reliability metric is absolute and directly calculates the probability that the difference between the model prediction and the experimental observation is within an acceptable tolerance limit.

4.5 Summary

Modeling and simulation is an integral part of many engineering applications. These models are used in different stages of the life-cycle of engineering systems. The system is not only analyzed but also designed using computational models; further, during the operations stage, models are used to monitor the system health, and therefore aid in diagnosis and prognosis.

The development of a computational model is accompanied by several sources of errors and uncertainty, and these sources of uncertainty, in turn, affect the quality of

the model prediction. Therefore, in order to quantify the uncertainty in the system-level prediction, it is not only necessary to quantify the input data uncertainty, but also essential to quantify the model uncertainty. While Chapter III focused on methods to address input data uncertainty, the current chapter addressed the issue of model uncertainty.

The quantification of model uncertainty is performed through three different activities, namely model verification, calibration and validation. These activities are not necessarily performed in a particular order, though it is preferable to perform model verification before calibration and/or validation. Calibration and validation activities may be performed in any order or even iteratively. The various aspects of verification, calibration, and validation, along with the contributions of this dissertation, are summarized below:

1. **Model Verification**: When a computational code is built to solve a mathematical equation, the true solution of the equation may not be estimated due to the presence of numerical solution approximations, rounding off errors, etc. The process of model verification quantifies the error between the true mathematical solution and the computer code output. Section 4.2 addressed the issue of model verification. Two different types of errors were considered: (1) discretization error due to the use of finite elements in computer codes; (2) solution approximation through surrogate models. While the former is a deterministic error, the latter is stochastic. These errors need to be included in uncertainty propagation; this will be demonstrated in the forthcoming chapters.

2. **Model Calibration**: Physics-based models have parameters which need to be calibrated/estimated so that the model predictions are in better agreement with the experimental observations. Section 4.3 discussed least squares-based,

likelihood-based and Bayesian methods for model parameter calibration. Section 4.3.8 developed Bayesian methods to perform model calibration under uncertainty, including interval data, partially characterized data, etc. It was observed that the Bayesian approach provides a systematic procedure to account for these types of uncertainty during the calibration procedure. The proposed methodology was illustrated by calibrating the non-linear stiffness term in a model used to predict the energy dissipation due to friction at a lap joint in a mechanical component.

3. **Model Validation**: Model validation quantifies the extent of agreement between the model prediction and the reality that the model is intended to represent. Section 4.4 investigated two approaches - Bayesian hypothesis testing and model reliability method - for model validation under uncertainty, including interval data, uncharacterized data, and time series data. The two methods are based on two different philosophies and hence, their results cannot be compared with each other. Further, there are several challenges in both these methods; in the former case, the choice of the alternate PDF is subjective, and in the latter case, the model reliability metric may yield low reliability values as the data size increases. Future work needs to address these issues and propose solutions in order to overcome the challenges with respect to the implementation of both these methods.

One challenge, especially when the system response prediction is based on multiple component-level and subsystem-level models, is that these uncertainty quantification activities need to be performed for each of the lower-level models. Therefore, in order to quantify the uncertainty in the system-level response, it is necessary to integrate the results from the uncertainty quantification activities performed at multiple levels. The development of a computational methodology for such integration is the focus of

Chapter VI. Prior to Chapter VI, Chapter V discusses the application of the proposed methods for data uncertainty quantification and model uncertainty quantification to fatigue crack growth analysis, as a case study.

CHAPTER V

CASE STUDY: FATIGUE CRACK GROWTH ANALYSIS

5.1 Introduction

The process of fatigue crack growth is affected by many sources of variability, such as loading, material properties, geometry and boundary conditions. Therefore, it is appropriate to describe the crack size after a certain number of load cycles through a probability distribution. Probabilistic fracture mechanics is an extensive area of research and numerous studies have addressed both model-based [151–172] and data driven [173–176] techniques for probabilistic crack growth and life prediction. This chapter focuses on model-based methods which include both probabilistic techniques [151, 152] and advanced computational mechanics techniques [153–155].

Probabilistic crack growth analysis has been applied to both metals (e.g. Johnson and Cook [156], Maymon [157]) and composites [158–162]. Practical applications of these methods include nuclear structural components [163], helicopter gears [164], gas turbine engines [165], and aircraft components [166, 167]. These developments have led to software for probabilistic fracture mechanics [168], and several commercial software tools such as DARWIN [169], GENOA [170], and other software tools that build probabilistic analysis around well-established codes such as AFGROW [171] and FASTRAN [172].

The aforementioned studies on probabilistic crack growth prediction have primarily included only the effects of natural variability in loading conditions, material and geometric properties, and crack growth model parameters. However, many of them have not explicitly accounted for other sources of uncertainty such as data uncertainty and other types of model uncertainty and error. The different types of error

and uncertainty appear at different stages of analysis, and may combine in nonlinear, nested, or iterative manner. Further, some of the errors are deterministic (e.g., finite element discretization error) while some others are stochastic (e.g., material properties). It is essential to systematically account for all these sources of error and uncertainty in order to ensure robustness in model calibration, model validation, and uncertainty quantification. This chapter includes the different sources of variability, data uncertainty, model uncertainty and errors that affect fatigue crack growth analysis, and develops a methodology to quantify the uncertainty and assess the validity in the overall crack growth model prediction (or prognosis).

Saxena et al. [177] reviewed several metrics to assess the performance of prognosis algorithms. These metrics are based on (1) observed error, (2) standard deviation of the observed quantity, (3) sensitivity, (4) reliability, and (5) cost-benefit analysis. These metrics do not directly quantify the confidence in the model prediction and do not delineate the contributions of the various sources of uncertainty. As explained earlier in Section 4.4 in Chapter IV, a rigorous approach to model validation needs to explicitly account for the different sources of uncertainty and quantitatively judge the performance of the model [103]. This chapter uses the Bayesian hypothesis testing methodology developed in Section 4.4.1 in order to quantify the confidence in the fatigue crack growth prediction; recall that this method quantifies the extent to which the data supports the model and directly computes the probability of the model being correct.

In this chapter, the various quantities of interest and different types of uncertainty (physical variability, data uncertainty, and model uncertainty/errors) are connected through a Bayesian network. Further, for use in time-dependent problems, a dynamic Bayesian network is constructed and this can be useful in two ways: (1) in an inverse problem [178] where the uncertainty in the model parameters is quantified using

experimental evidence; and (2) in a forward problem [179] that aids in probabilistic prediction.

In the inverse problem of calibration for fatigue crack growth prediction, previous techniques have been used to calibrate either a single parameter [180, 181] or the parameters of a single model [182]. Further, these studies [180–182] have not explicitly included the effect of different types of uncertainty - physical variability, data uncertainty, and model uncertainty and errors - on calibration. As a result, all the difference between the model prediction and experimental observation gets attributed to the calibration parameters without acknowledging or accounting for the other sources of error and uncertainty. In this chapter, the effect of the three different types of uncertainty (physical variability, data uncertainty, and model uncertainty and errors) are explicitly included and multiple parameters and model errors are calibrated simultaneously, using the calibration technique discussed earlier in Section 4.3 in Chapter IV.

The forward problem involves the use of the calibrated model for probabilistic crack growth prediction. Probabilistic damage prediction aids in effective prognosis under uncertainty, and helps to forecast the remaining useful life of the structural/mechanical system under study. In this chapter, a sampling-based methodology is proposed to include all three types of uncertainty in this prediction; then this prediction is validated against validation data collected through inspection.

Different types of data situations are considered for both calibration and validation - (1) data available for constructing the probability distributions of input variables may be sparse, leading to uncertainty regarding their distribution types and parameters; (2) inspection data used for calibration and/or validation may be of three different types: crack not detected, crack detected but size not measured, or crack detected with size measurement; and (3) measurement errors in inspection data.

Further, different models - long crack growth law, wheeler's retardation model, characteristic plane approach, finite element analysis to calculate stress intensity factor, and surrogate model to replace the finite element analysis - are integrated for the purpose of fatigue crack growth prediction. The presence of several models must not be confused with hierarchical system models discussed earlier in Section 1.3.1. Hierarchical system models refer to multiple component-level, subsystem-level and system-level models, and data could be available at multiple levels. In this chapter, the goal is to perform uncertainty quantification and validation in a single level, with respect to a single engineering component subjected to fatigue crack growth.

The steps of the proposed methodology can be summarized as follows: (1) connect different models (finite element model, surrogate model, crack growth law, retardation model, and characteristic plane model) through a Bayesian network; (2) quantify the errors in each model, treating deterministic and stochastic errors differently; (3) include the different sources of uncertainty physical variability, data uncertainty, and model uncertainty in the Bayesian network; (4) use global sensitivity analysis to select calibration parameters; (5) collect inspection data and update the parameters in multiple models; (6) include different cases of inspection such as crack not detected, crack detected but size not measured, and crack detected with size measurement; and (7) use a Bayesian metric to assess the validity of the model prediction. (Note that the data for model calibration and model validation should be different and independent.)

The above steps are implemented in this chapter for planar crack growth analysis, in the context of linear elastic fracture mechanics (LEFM) with small scale plasticity. Different types of model uncertainty and errors crack growth model uncertainty, surrogate model uncertainty and finite element discretization error - are considered explicitly. Deterministic errors are addressed by correcting where they arise and stochastic errors are included through sampling. The effects of model assumptions

such as planar crack, LEFM, equivalent stress intensity factor for multi-modal fracture, and crack retardation model are not individually quantified. However, in the model validation step, the difference between the model prediction and the experimental observation obviously includes the contribution of these sources.

The rest of the chapter is organized as follows. Section 5.2 discusses the crack growth modeling procedure. Section 5.3 discusses several sources of uncertainty and proposes methods to handle them.Section 5.4 discusses global sensitivity analysis to screen the model parameters for calibration, and outlines the proposed Bayesian inference technique for calibrating these parameters.Section 5.5 extends Bayesian hypothesis testing to time-dependent problems, in order to assess the validity and the confidence in the crack growth model prediction.Section 5.6 illustrates the proposed methods using a numerical example, surface cracking in a cylindrical component.

5.2 Crack Growth Modeling

Different models such as the modified Paris' law (for long crack growth analysis), Wheeler's retardation model (for variable amplitude loading), characteristic plane approach (to calculate an equivalent stress intensity factor in the presence of multi-axial loading) are connected together to predict the crack growth as a function of number of load cycles. Note that these models are used only for the purpose of illustration and any other combination of appropriate models may be used. The focus of this chapter is on uncertainty quantification and model validation and these techniques are applicable to different crack growth analysis procedures.

Consider any long crack growth law used to describe the relationship between da/dN and ΔK, where N represents the number of cycles, a represents the crack size and ΔK represents the stress intensity factor range. This dissertation uses a modified Paris' law [183] for illustration purposes and includes Wheelers retardation

model [184] as:

$$\frac{da}{dN} = \phi^r C(\Delta K)^n \Big(1 - \frac{\Delta K_{th}}{\Delta K}\Big)^m \tag{5.1}$$

In Eq. 5.1, ΔK_{th} refers to the threshold stress intensity factor range and ϕ^r refers to the retardation parameter which can be defined as:

$$\phi^r = \begin{cases} \Big(\frac{r_{p,i}}{a_{OL}+r_{p,OL}-a_i}\Big)^{\lambda} & \text{if } a_i + r_{p,i} < a_{OL} + r_{p,OL} \\ 1 & \text{if } a_i + r_{p,i} \geq a_{OL} + r_{p,OL} \end{cases} \tag{5.2}$$

In Eq. 5.2, a_{OL} is the crack length at which the overload is applied, a_i is the current crack length, $r_{p,OL}$ is the size of the plastic zone produced by the overload at a_{OL}, $r_{p,i}$ is the size of the plastic zone produced at the current crack length a_i, and λ is the curve fitting parameter for the original Wheeler model termed the shape exponent [185]. Sheu et al. [186] and Song. et al. [187] observed that crack growth retardation actually takes place within an effective plastic zone. Hence the size of the plastic zone can be calculated in terms of the applied stress intensity factor (K) and yield strength (σ) as:

$$r_p = \alpha \Big(\frac{K}{\sigma}\Big)^2 \tag{5.3}$$

where α is known as the effective plastic zone size constant which is obtained experimentally [186]. Eq. 5.2 and Eq. 5.3 can be used in combination with Eq. 5.1 under the assumption of small-scale plasticity, where the plastic zone size is estimated using linear-elastic fracture mechanics (LEFM).

The expressions in Eq. 5.2 and Eq. 5.3 can be combined with Eq. 5.1 and used to calculate the crack growth as a function of number of cycles. In each cycle, the stress intensity factor can be expressed as a function of the crack size (a), loading (L) and

angle of orientation (χ). Hence, the crack growth law in Eq. 5.1 can be rewritten as

$$\frac{da}{dN} = g(a, L, \chi) \tag{5.4}$$

Note that the long crack growth model is not applicable to the short crack growth regime and the concept of an equivalent initial flaw size (EIFS) was proposed to bypass short crack growth analysis and make direct use of a long crack growth law for fatigue life prediction. The equivalent initial flaw size θ, i.e., the initial condition of the differential equation in Eq. 5.1, can be calculated from material properties (ΔK_{th}, the threshold stress intensity factor and $\Delta\sigma_f$, the fatigue limit) and geometric shape factor (Y) as derived by Liu and Mahadevan [188]:

$$\theta = \frac{1}{\pi}\Big(\frac{\Delta K_{th}}{Y\Delta\sigma_f}\Big)^2 \tag{5.5}$$

By integrating the expression in Eq. 5.1 starting from θ, the number of cycles (N) to reach a particular crack size a_N can be calculated as:

$$N = \int dN = \int_{\theta}^{a_N} \frac{1}{\phi^r C(\Delta K)^n \big(1 - \frac{\Delta K_{th}}{\Delta K}\big)^m} da \tag{5.6}$$

The stress intensity factor range ΔK in Eq. 5.6 can be expressed as a closed form function of the crack size for specimens with simple geometry subjected to constant amplitude loading. However, this is not the case in many mechanical components, where ΔK depends on the loading conditions, geometry and the crack size. Further, if the loading is multi-axial (for example, simultaneous tension, torsion and bending), then the stress intensity factors corresponding to three modes need to be taken into account. This can be accomplished using an equivalent stress intensity factor. If K_I, K_{II}, K_{III} represent the mode-I, mode-II and mode-III stress intensity factors respectively, then the equivalent stress intensity factor K_{eqv} can be calculated using

a characteristic plane approach proposed by Liu and Mahadevan [189] as:

$$K_{eqv} = \frac{1}{B}\sqrt{(K_I)^2 + (\frac{K_{II}}{s})^2 + (\frac{K_{III}}{s})^2 + A(\frac{K_H}{s})^2} \tag{5.7}$$

In Eq. 5.7, K_H is related to hydrostatic stress, and s is the ratio of K_{II} and K_I evaluated at a specific crack growth rate (da/dN). A and B are material parameters. The characteristic plane approach is applicable only when the crack surface can be approximated to be planar. The use of the characteristic plane approach for crack growth prediction under multi-axial variable amplitude loading has been applied to cracks in railroad wheels [189] and validated earlier with several data sets [190, 191].

Each cycle in the integration of Eq. 5.6 involves the computation of ΔK_{eqv} using a finite element analysis owing to (1) complicated geometry, and (2) variable amplitude, multi-axial loading. Repeated evaluation of this finite element analysis renders this integration extremely expensive. Hence, it is computationally more affordable to substitute the finite element model with an inexpensive surrogate model (also known as response surface model). Different kinds of surrogate models (polynomial chaos [45], support vector regression [46], relevance vector regression [47], and Gaussian Process interpolation [51, 52] have been explored and the Gaussian process (GP) modeling technique has been found to have the least error for the purpose of predicting ΔK_{eqv} [181] . A few runs of the finite element analysis are used to train the GP surrogate model and then, this GP model is used to predict the stress intensity factor for other crack sizes and loading cases (for which finite element analysis has not been carried out).

The details of the construction of the Gaussian process were discussed earlier in Section 2.8. Further, the construction of the Gaussian process surrogate model in order to predict the equivalent stress intensity factor has been documented [179, 181]. This equivalent stress intensity factor is then used in cycle-by-cycle integration of the

crack growth law, thereby calculating the crack size (A) as a function of number of load cycles (N). The entire procedure for the adopted crack growth analysis is summarized in Fig. 5.1.

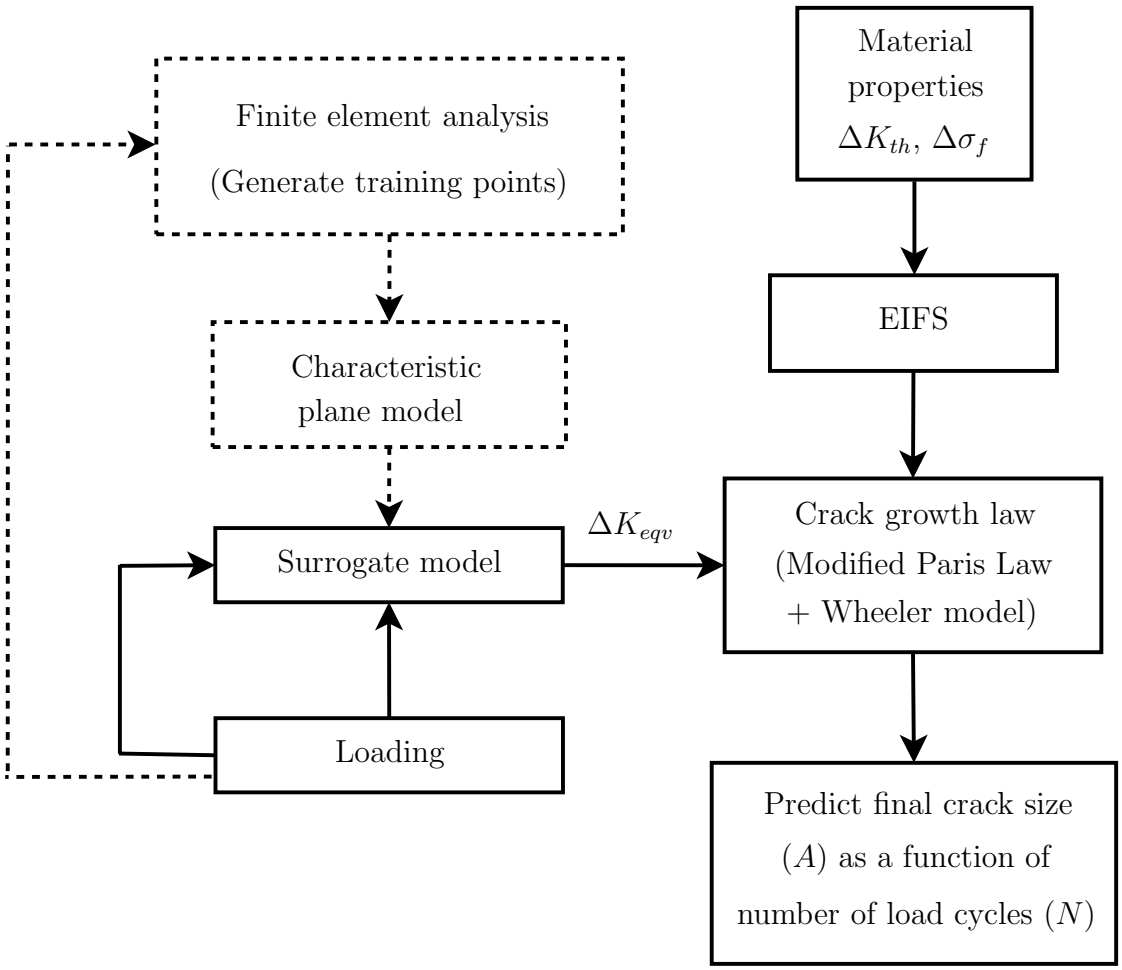


Figure 5.1: Deterministic Crack Propagation Analysis

In Fig. 5.1, note that the finite element analysis and the construction of surrogate model are performed offline, i.e. before the start of crack growth analysis. Crack propagation analysis is done only with the surrogate model.

The algorithm shown in Fig. 5.1 for crack propagation analysis is deterministic and does not account for errors and sources of uncertainty. The next section discusses different sources of uncertainty associated with each of the blocks in Fig. 5.1.

5.3 Sources of Uncertainty

This section proposes methods to include the different sources of uncertainty in the crack growth analysis algorithm shown in Fig. 5.1. These sources of uncertainty can be classified into three different types – physical variability, data uncertainty and model uncertainty – as shown in Fig. 5.2.

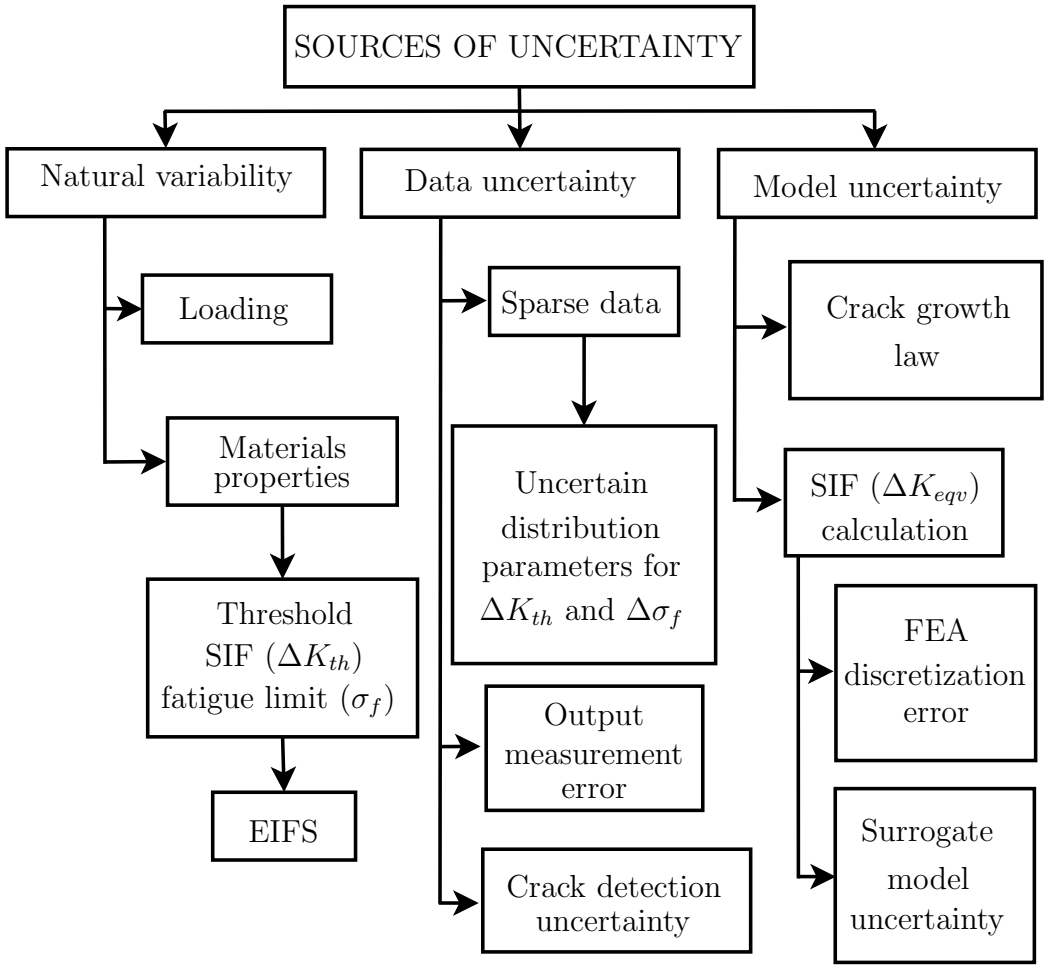


Figure 5.2: Sources of Uncertainty in Crack Growth Prediction

Fig. 5.2 shows the different sources of error and uncertainty considered in this chapter for the sake of illustration of the proposed methodology. There are several other sources of uncertainty that are not considered here. The physical variability in geometry, boundary conditions and some material properties (modulus of elasticity,

friction coefficient, Poisson ratio, yield stress, ultimate stress, etc.) is assumed to less significant and hence these quantities are treated to be deterministic. However, these can be included in the proposed framework by constructing different FEA models (for different geometry, boundary conditions, and material properties) and use multiple runs of these FEA models to train the Gaussian process surrogate model. Then, these parameters can also be treated as inputs to the surrogate model and sampled randomly in the uncertainty quantification procedure explained later in Sections 5.4 and 5.5. Further, there are several other types of model errors that arise due to (1) the assumption of a planar crack; (2) the assumption of linear elastic behavior and use of LEFM; (3) the use of a characteristic plane approach (to calculate an equivalent stress intensity factor) and parameters of the characteristic plane model in Eq. 5.7; (4) errors in the estimation of retardation coefficient fr in Wheeler's model, etc. These errors are not considered in this study and the quantification of these errors is not trivial; these errors will be considered in future work. The focus of this chapter is not the quantification of all the error sources but to develop a methodology to assess the validity of the crack growth prediction by systematically accounting for the various sources of uncertainty and errors. The error and uncertainty terms considered in this study adequately illustrate the various techniques for model calibration, validation and uncertainty quantification.

The following subsections discuss each source of uncertainty and propose methods to handle them.

5.3.1 Physical or Natural Variability

This case study considers the physical variability in (i) loading conditions, and (ii) material properties (threshold stress intensity factor ΔK_{th} and fatigue limit $\Delta\sigma_f$).

As mentioned earlier, the variability in other material properties such as modulus of elasticity, Poisson ratio, etc. is not considered.

5.3.2 Data Uncertainty

Two types of data uncertainty are addressed in this case study. First, sparse data is assumed to be available for variability characterization, leading to distribution parameter uncertainty in the distributions of material properties (ΔK_{th} and $\Delta\sigma_f$). The total probability-based integration approach proposed in Section 3.4 in Chapter III is used for data uncertainty quantification in this case study.

Second, field inspections are used to collect crack growth data for calibration and validation; the presence of crack detection uncertainty and measurement errors are also included in this case study. For an in-service component, non-destructive inspection (NDI) is commonly used for damage detection. Several metrics could be used to evaluate the performance of NDI, such as probability of detection (POD), flaw size measurement accuracy, and false call probability (FCP). These criteria are developed from different methods, and they are used to evaluate different aspects of NDI performance. However, Zhang and Mahadevan [192] showed that these quantities can be mathematically related. POD and FCP can be derived from size measurement accuracy, which measures the difference between actual values and observed values of the crack size.

In the context of calibration and/or validation, inspection results may be of three different types: (a) crack is not detected; (b) crack is detected but size not measured; and (c) crack is detected and size is measured. The concept of POD can be employed for the first two cases and for the case of detecting a crack and also measuring its size, size measurement accuracy can be employed. Based on the above consideration, size measurement accuracy can be used to quantify the uncertainty in experimental crack

growth data, with the following expression determined by regression analysis [192]:

$$a_m = \beta_0 + A \times \beta_1 + \epsilon_m \tag{5.8}$$

In Eq. 5.8, is the measured flaw size; A is the actual flaw size; β_0 and β_1 are the regression coefficients; ϵ_m represents the unbiased measurement error, assumed as a normal random variable with zero mean and standard deviation σ_ϵ. The value of σ_ϵ is different for each inspection technique used.

5.3.3 Model Uncertainty and Errors

The proposed methodology uses several models (finite element model, characteristic plane model, surrogate model, crack growth model, retardation model, etc.), and each of these models has its own error/uncertainty. In the discussion below, all the errors except the error due to the use of the characteristic plane approach and the uncertainty in the calculation of the retardation coefficient are considered. Some of them are deterministic while others are stochastic; the two types of errors need to be treated in different ways. In the discussion below, the quantification of finite element discretization and surrogate model error are solution approximation errors associated to the calculation of stress intensity factor, and hence grouped together.

5.3.3.1 Uncertainty in Crack Growth Model

More than 20 different crack growth laws (e.g., Paris law, Foreman's equation, Weertman's equation) have been proposed in the literature. The mere presence of many such different models explains that none of these models can be applied universally to all fatigue crack growth problems. Each of these models has its own limitations and uncertainty. In this dissertation, a modified Paris law has been used

only for illustration; however, the proposed methodology for uncertainty quantification and model validation can be implemented using any crack growth model.

The uncertainty in crack growth model can be subdivided into two different types: crack growth model error and uncertainty in model coefficients. If ϵ_{cg} is used to denote the crack growth model fitting error, then the modified Paris law (including the Wheeler retardation term) can be rewritten as in Eq. 5.9. Note that this error term is assumed to represent the difference between the model prediction and the experimental observation. Conventionally, in statistical model fitting, a normal error term is added to the model. This study uses a lognormal and hence multiplicative error term as shown in Eq. 5.9.

$$\frac{da}{dN} = \phi^r C(\Delta K)^n \Big(1 - \frac{\Delta K_{th}}{\Delta K}\Big)^m \epsilon_{cg} \tag{5.9}$$

An estimate of the variance of ϵ_{cg} can be obtained while calibrating the model parameters based on the available inspection data. The model coefficients in Paris law are C, m and n, and the uncertainty in these parameters can be represented through probability distributions.

5.3.3.2 Solution Approximation Errors

Several finite element runs for some combination of input-output variable values are used to train the Gaussian process surrogate model in order to predict the equivalent stress intensity factor. The input variables are loading and crack geometry. Once the surrogate model is constructed, it can be used to evaluate the stress intensity factor for other (untrained) combinations of input variable values.

There are two errors – discretization error and surrogate model error – that need to be quantified in this procedure. The quantification of these errors were discussed earlier in Sections 4.2.1 and 4.2.2 respectively. First, the discretization error is used to

correct the finite element solution, and the corrected values are in turn used to train the surrogate model. The prediction of the surrogate model at a particular input realization is a Gaussian random variable, and the surrogate model uncertainty can be included by drawing random samples based on the predicted mean and variance.

5.4 Calibration of Model Parameters

This section explains the Bayesian calibration technique used to infer the probability distributions of the model parameters and modeling errors using experimental data. The calibration parameters are assigned prior distributions, and these distributions are updated after collecting experimental evidence corresponding to the model output (crack inspection after a particular number of loading cycles).

5.4.1 Bayesian Network

There are several possible quantities that can be updated using the given procedure. These include: (1) equivalent initial flaw size; (2) parameters of modified Paris's law (C, n, m); (3) error (ϵ_{cg}) of the modified Paris law; (4) material properties such as threshold stress intensity factor (ΔK_{th}), and fatigue limit ($\Delta\sigma_f$); (5) variance of the output (crack size after N cycles) measurement error (ϵ_m), etc. All of these parameters can be connected in a Bayesian network as shown in Fig. 5.3.

Note that this is a dynamic Bayesian network, which connects the variables in one load cycle (i) to the next load cycle ($i+1$).

Though it is theoretically possible to calibrate all the parameters simultaneously, a multi-dimensional calibration procedure is computationally intensive. Also, in several situations, only a few parameters contribute effectively to the overall uncertainty in the final model prediction. It is advantageous to identify only such "important" parameters and calibrate them. Global sensitivity analysis is used to quantify the

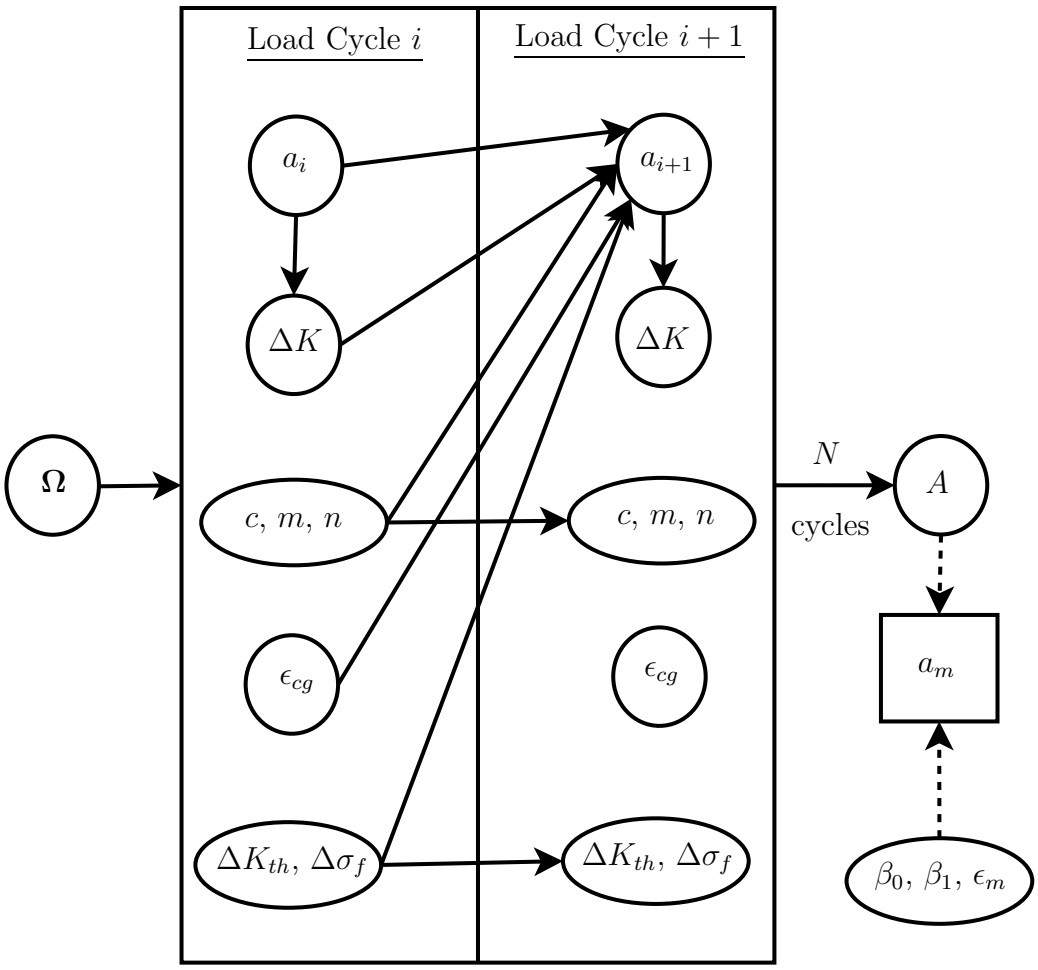


Figure 5.3: Dynamic Bayesian Network: Crack Growth Analysis

contribution of the various sources of uncertainty to the overall uncertainty in the crack growth prediction, and uses the results of global sensitivity analysis to identify the parameters that need to be calibrated.

5.4.2 Global Sensitivity Analysis

The method of global sensitivity analysis has previously been applied only to naturally varying random inputs; this method is extended to include data uncertainty and model uncertainty as well. Since all the quantities have been represented using probability distributions, the method of global sensitivity analysis discussed in Section 2.6

can be used to analyze the contributions of physical variability, data uncertainty and model uncertainty to the overall uncertainty in crack size prediction. Note that the equivalent initial flaw size has physical variability with distribution parameter uncertainty; hence, the method developed in Section 3.5 can be used to quantify the contributions of both the physical variability and distribution parameters to the crack size uncertainty.

Since the crack size is predicted as a function of number of load cycles, the uncertainty in crack size is also computed as function of number of load cycles. Therefore, the sensitivities of the various sources of uncertainty are also computed as a function of number of load cycles.

The quantities that have a high influence on the variance of the final crack growth prediction are alone selected for calibration. While this screening procedure is based on second-moment calculation, the model calibration and validation techniques explained in the following sections consider the entire probability distribution for all quantities.

5.4.3 Model Calibration

Let $\mathbf{\Omega}$ denote the vector of quantities that are selected for calibration. Assume that there is a set of m experimental data points, i.e. crack inspection after N loading cycles. Each inspection may produce three possible results: (1) no crack detected; (2) crack detected but size not measured; and (3) crack detected and size measured to be A. This is the first set of data, referred to as the calibration data.

Bayesian updating is a three step procedure:

1. Prior probability distributions are assumed for each of the parameters. The prior distribution of EIFS is calculated using Eq. 5.5. The crack growth model

error (ϵ_{cg}) has zero mean and the standard deviation is assumed to have a non-informative prior. The prior distributions of crack growth law parameters are obtained from Liu and Mahadevan [188].

2. Construction the likelihood function, i.e. $L(\mathbf{\Omega})$ by including the different types of uncertainty, using the methods developed in Section 4.3.8.

3. Calculate the posterior PDF $f_{\mathbf{\Omega}}(\mathbf{\Omega}|D)$ using Bayesian inference as explained in Section 2.4. The joint PDF of $\mathbf{\Omega}$ is used to calculate the marginal distributions of the individual parameters.

In step 2, the likelihood of $\mathbf{\Omega}$ is calculated as the probability of observing the data conditioned on Ω and hence is a function of $\mathbf{\Omega}$. For every given $\mathbf{\Omega}$, a Monte Carlo analysis is required for the calculation of likelihood along the following steps:

1. Construct the Gaussian process surrogate model as explained in Section 5.2. Include the sources of uncertainty as explained in Section 5.3.

   For a given $\mathbf{\Omega}$:

2. Generate a loading history (N_i cycles) as explained in Section 5.3.

3. Use the deterministic prognosis methodology in Section 5.2 to calculate the final crack size at the end of N_i cycles.

4. Repeat steps II and III and calculate the probability distribution of crack size at the end of N_i (for $i = 1$ to m) cycles. Let this distribution be denoted by $f(a)$. Use Eq. 5.8 to calculate $f(a_m|a)$. This probability density function can be used to calculate the likelihood of $\mathbf{\Omega}$.

In order to calculate the likelihood function $L(\mathbf{\Omega})$, calibration data needs to be collected using inspection. For a particular inspection after N_i cycles, there are three

possible cases:

Case 1: No crack detected.

$$L(\mathbf{\Omega}) \propto 1 - \int \int f(a_m|a) f(a|N_i, \mathbf{\Omega}) da da_m \tag{5.10}$$

Case 2: Crack is detected but size not measured.

$$L(\mathbf{\Omega}) \propto \int \int f(a_m|a) f(a|N_i, \mathbf{\Omega}) da da_m \tag{5.11}$$

Case 3: Crack is detected and the size is measured to be A_i.

$$L(\mathbf{\Omega}) \propto \int f(a_m = A_i|a) f(a|N_i, \mathbf{\Omega}) da \tag{5.12}$$

Note that a_m and a vary in the interval (0, ∞).

If different inspections give any of the three types of crack information, the corresponding likelihood function for each data point is calculated using Eq. 5.10 - 5.12 and then all the likelihood functions are multiplied to get the overall likelihood function for the entire inspection data set. Finally, the likelihood is multiplied with the prior and normalized to calculate the joint posterior distribution of $\mathbf{\Omega}$ [52], and this joint distribution can be marginalized to calculate the individual distributions [27].

5.5 Model Validation

This section uses the Bayesian hypothesis procedure in order to assess the validity and the confidence in the prediction of the fatigue crack growth models. Consider the crack-growth algorithm discussed in Section 5.2. The probability distribution of the crack size (a) can be calculated as a function of number of load cycles (N) after accounting for the various sources of uncertainty in a systematic manner using the

methods in Sections 5.2 and 5.3. Let $f(a|N)$ denote the corresponding probability density function. The aim of this section is to assess the validity of the underlying crack growth model and propose a metric to quantify the confidence in the model prediction. It is assumed that a second set of data (D), referred to as validation data, is available for this purpose, and is independent of the calibration data. Recall that the Bayes factor validation metric can be computed as:

$$B = \frac{P(D|H_0)}{P(D|H_1)} \tag{5.13}$$

In Eq. 5.13. The calculation of the numerator is based on the available data (D), and the prediction of the crack size $f(a|N)$ and $f(a_m|a)$ is calculated based on Eq. 5.8. For a particular inspection after N_i cycles, there are three possible cases:

Case 1: No crack detected.

$$L(\mathbf{\Omega}) \propto 1 - \int\int f(a_m|a)f(a|N_i)dada_m \tag{5.14}$$

Case 2: Crack is detected but size not measured.

$$L(\mathbf{\Omega}) \propto \int\int f(a_m|a)f(a|N_i)dada_m \tag{5.15}$$

Case 3: Crack is detected and the size is measured to be A_i.

$$L(\mathbf{\Omega}) \propto \int f(a_m = A_i|a)f(a|N_i)da \tag{5.16}$$

Note that Eq. 5.14 - 5.16 are different from Eq. 5.10 - 5.12. Previously, $f(a|N, \mathbf{\Omega})$ was calculated for a given realization of calibration parameters $\mathbf{\Omega}$. This is not the case in Eq. 5.14 - 5.16 where the uncertainty in the parameters $\mathbf{\Omega}$ (sampled using posterior probability distributions of $\mathbf{\Omega}$) is included in $f(a|N)$.

The calculation of the denominator is related to the hypothesis that the model is not correct. Hence, it is necessary to assume a distribution for $f(a|N)$ under this hypothesis. As there is no information available for this hypothesis, a non-informative uniform prior distribution for $f(a|N)$ is used and the denominator term $P(D|H_1)$ is calculated based on Eq. 5.14 - 5.16. Then the Bayes factor can be calculated using the likelihood ratio in Eq. 5.13.

There may be multiple inspections at multiple time instants for the sake of crack growth validation. If these multiple inspections are performed are multiple components, the inspections can be assumed to be independent of each other, the overall Bayes factor (that accounts for data from multiple time instants) will be equal to the product of the individual Bayes factors. Assume that the data D is composed of data from several independent inspections D_1, D_2, D_3....D_n and the Bayes factor corresponding to data D_i is B_i. Then the overall Bayes factor is the product of all B_i's. On the other hand, if the multiple inspections are performed on the same specimen, then the method developed in Section 4.4.1.4 can be used to compute the overall Bayes factor (B). The confidence associated with the prediction is thus calculated as $\frac{B}{B+1}$.

In summary, Sections 5.4 and 5.5 proposed a Bayesian methodology for (1) calibration and uncertainty quantification of model parameters using inspection data (crack size measurements after a number of cycles, including cases where no cracks were detected, and crack detections without measurements), and (2) assessing the validity of the crack growth prediction by comparing against a second independent set of validation data. The following section illustrates the proposed methodology using a numerical example.

5.6 Numerical Example

A two-radius hollow cylinder with an elliptical surface crack in fillet radius region is chosen for illustrating the proposed methodology, as shown in Fig. 5.4.

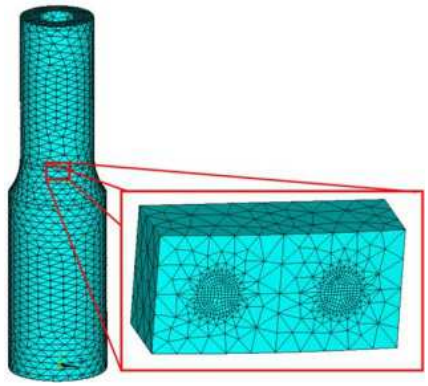


Figure 5.4: Surface Crack in Cylindrical Component

The commercial finite element software ANSYS [193] is used to build and analyze the finite element model. A sub-modeling technique is used near the region of the crack for accurate calculation of the stress intensity factor. As mentioned earlier in Section 5.3, the material and geometric properties of the specimen are assumed to be deterministic and are listed in Table 5.1 and Table 5.2 respectively.

Table 5.1: Material Properties

Aluminium 7075-T6	
Property	**Value**
Modulus of elasticity	72 GPa
Poisson ratio	0.32
Yield stress	450 MPa
Ultimate stress	510 MPa

Table 5.2: Geometric Properties

Cylinder Properties	
Property	**Value**
Length	152.4 mm
Inside radius	8.76 mm
Outside radius (Narrow)	14.43 mm
Outside radius (Wide)	17.78 mm

A variable amplitude multi-axial (bending and torsion) loading is considered in this study. A block loading history is considered (sample shown in Fig. 5.5), for the purpose of illustration.

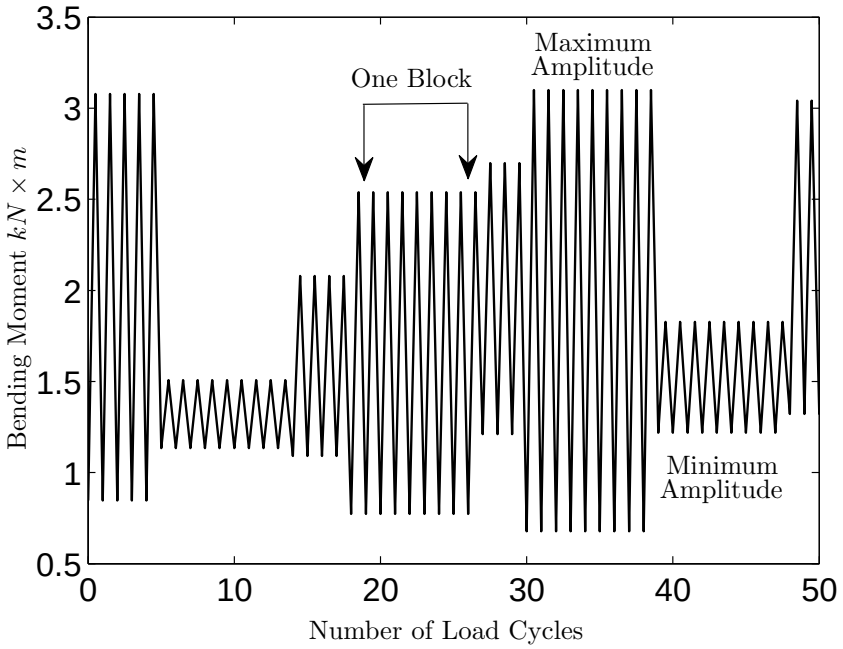


Figure 5.5: Sample Block Loading History

The block length is assumed to be a random variable and the maximum and minimum amplitudes in each block are also treated as random variables. To generate each block of loading, first the block length is randomly sampled and then the maximum and minimum amplitudes for that block are randomly sampled. The entire loading history is generated by repeating this process and creating several successive blocks to occupy the duration of interest.[1]

All the different types of uncertainty in Fig. 5.2 are included in this numerical example, as shown in Table 5.3.

This finite element model is run for 10 different crack sizes and 6 different loading cases and these results are used to train the Gaussian process surrogate model for the calculation of the stress intensity factor. In each run of the finite element analysis, the discretization error is calculated and the corrected solution in turn is used to train

[1]Ling et al. [194, 195] have extended this study to include other types of time-dependent models for loading, such as rainflow counting, Markov processes, ARIMA models, etc.

Table 5.3: Types of Uncertainty

Type	**Quantities**	**Details**
Natural Variability	Loading	Random Block Loading Block Length ˜ N(0, 500) Maximum Amplitude ˜ N(μ_{max}, σ_{max}) Minimum Amplitude ˜ N(μ_{max}, σ_{max}) $\mu_{max} \sim$ Uniform(1.355, 0.263) kNm $\sigma_{max} \sim$ Uniform(0.455, 0.130) kNm $\mu_{min} \sim$ Uniform(2.710, 0.260) kNm $\sigma_{min} \sim$ Uniform(0.455, 0.130) kNm
Natural Variability	Material Properties (for prior distribution of EIFS)	ΔK_{th} and $\Delta\sigma_f$ Distribution type, mean and standard deviation based on [188]
Data Uncertainty	Sparse data	Distribution parameters of ΔK_{th} and $\Delta\sigma_f$ are uncertain $\rightarrow$ Assumed CoV for mean = 0.1 & standard deviation = 1 $\rightarrow$ Use Eq. 3.8 to calculate PDFs of ΔK_{th} and $\Delta\sigma_f$ $\rightarrow$ Use Eq. 5.5 to calculate prior PDF of EIFS (θ)
Data Uncertainty	Inspection data (crack not detected, crack detected but size not measured, and crack size measured) considered in both calibration and validation	
Model Uncertainty	Crack Growth Model	Statistics of uncertain parameters based on [188]. Also see Table 5.4 for prior statistics
Model Uncertainty	Discretization Error	Calculated using Richardson Extrapolation Refer to Section 4.2.1
Model Uncertainty	Surrogate Model Error	Compute GP variance based on Eq. 2.19 and include in uncertainty propagation

the Gaussian process surrogate model. The maximum discretization error was found to be about 2% of the finest mesh solution. If the discretization errors were found to be large, then it indicates a coarsely resolved calculation, which may not be suitable in the context of model calibration.

Global sensitivity analysis was used to calculate the first order (individual) and total (combined) effects of the various crack growth parameters such as EIFS (θ), model parameters (C, m), crack growth model error (ϵ_{cg}), threshold stress intensity factor (ΔK_{th}), loading, etc. It was observed that the sensitivity of the parameters changed with time and hence, it was necessary to calculate the sensitivity indices as a function of number of cycles. After a significant number of loading cycles, the parameters EIFS θ, model parameter C, and crack growth model error (ϵ_{cg}) were

found to be the most significant terms. The first-order and total effects sensitivities of these quantities are shown in Figs. 5.6 - 5.8. Recall that the first-order effects describe the contribution of a particular variable to the uncertainty in the crack size, by itself, whereas the total effects describe the overall contribution of a particular variable to the uncertainty in the crack size, in combination with all other variables.

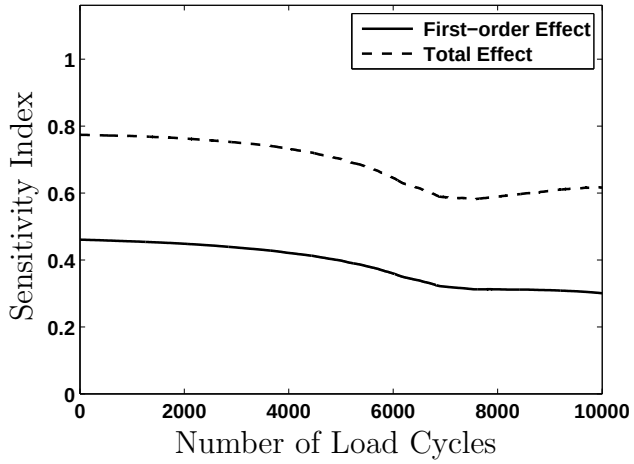


Figure 5.6: Variability of EIFS: Sensitivity Index

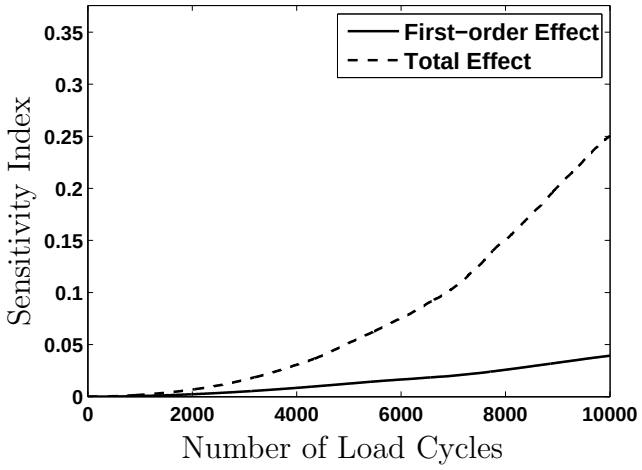


Figure 5.7: Crack Growth Parameter C: Sensitivity Index

The sensitivity indices of the other parameters were insignificant (less than 0.001) and hence only three parameters (EIFS θ, model parameter C, and crack growth

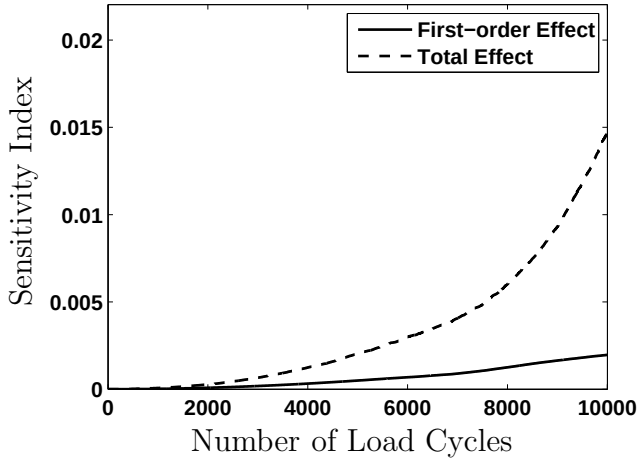


Figure 5.8: Crack Growth Error ϵ_{cg}: Sensitivity Index

model error ϵ_{cg}) are considered to have a dominant influence on the uncertainty in the crack growth prediction in comparison with the other parameters. Hence, these three are chosen to be the calibration parameters.

Experimental evidence of 100 data points (inspection) was simulated (based on assumed true distributions for the inference quantities) and used for calibration as explained in Section 5.4. 100 components were inspected at different time instants; no crack was detected in one component, the crack size was measured in 49 components, and the crack size was not measured in 50 components. This experimental data is used to calculate the statistics of the parameters using the likelihood function (Eq. 5.10 - 5.12) and Bayes theorem.

Note that Eq. 5.10 - 5.12 calculate the joint likelihood of the inference quantities (EIFS θ, model parameter C, and crack growth model error ϵ_{cg}) and hence, Bayes theorem gives the joint probability density function of these quantities. Hence, in addition to the marginal posterior distributions, the correlations between these quantities can also be estimated using the joint density function. For example, the marginal prior and posterior PDF of EIFS is shown in Fig. 5.9.

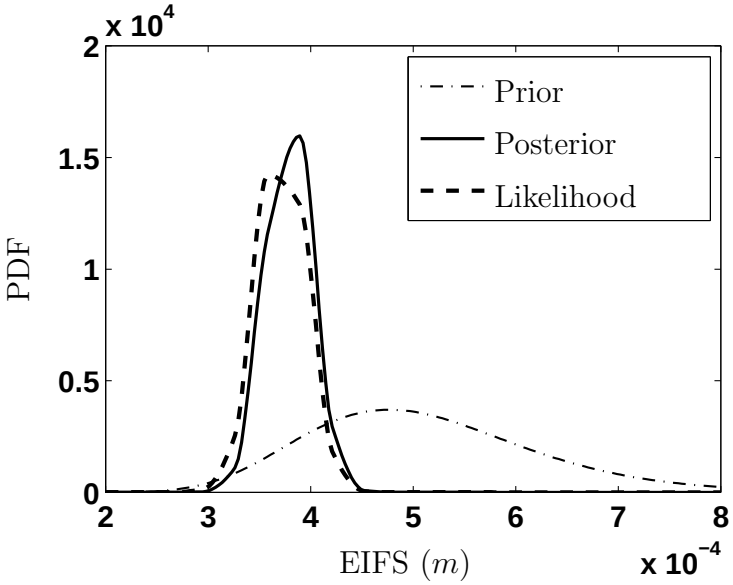


Figure 5.9: Prior and Posterior PDF of EIFS

The mean and the standard deviation of the prior distribution and the posterior distribution of all the three inferred quantities are given in Table 5.4.

Table 5.4: Results of Calibration

Quantity	EIFS θ (mm)	C (m/cycle)	$\epsilon_{cg}(mm/cycle)$
<u>True Statistics</u>			
Mean	0.3810	4×10^{-13}	1.0
Standard Deviation	0.0254	0.4×10^{-13}	0.05
<u>Prior Statistics</u>			
Mean	0.5126	6.5×10^{-13}	1.0
Standard Deviation	0.1146	4.0×10^{-13}	0.1
<u>Posterior Statistics</u>			
Mean	0.3788	5.7×10^{-13}	1.0
Standard Deviation	0.0222	3.1×10^{-13}	0.05

The Bayesian approach will be applicable even if the mean of the prior distribution is one or two orders of magnitude away from the truth as long as the uncertainty in the prior is large enough to contain the truth. In other words, the Bayesian approach will work if there is any overlap between the prior distribution and the truth.

(If there is no overlap, then the posterior PDF will be zero everywhere. Note

that the posterior distribution is proportional to the prior times the likelihood. In the "region of the truth", the prior will be zero as there is no overlap. Outside the "region of the truth", the likelihood will be zero.)

Having calibrated the parameters, the posterior distributions can be used to predict the distribution of crack size as a function of number of loading cycles. This PDF was denoted by $f(a|N)$ in Section 5.5. The mean, median and 95% probability bounds of the crack size prediction is shown in Fig. 5.10 as a function of number of cycles.

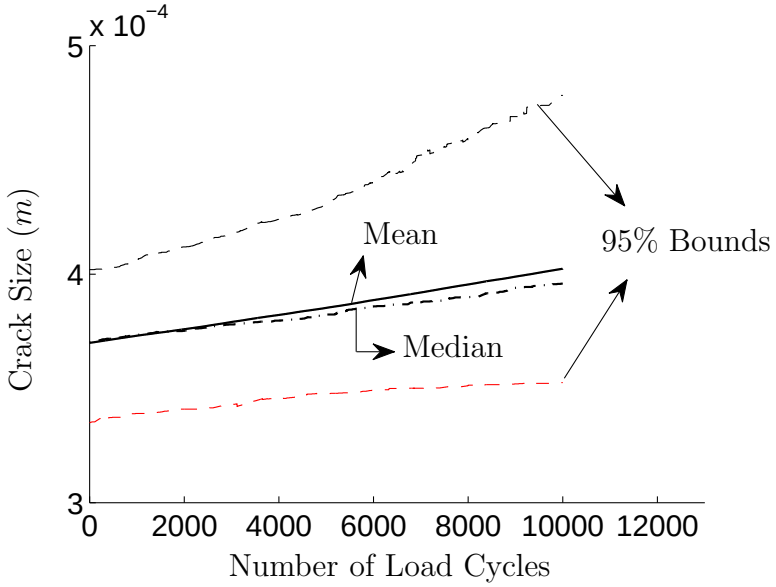


Figure 5.10: Uncertainty in Crack Growth Prediction

Additional independent validation data is assumed to be available and the Bayes factor is calculated for each data point individually. Recall that the Bayes factor is the ratio of the likelihoods of the null and the alternate hypothesis. Eq. 5.14 – 5.16 are used to calculate the $P(D|H_0)$ where H_0 represents the hypothesis that the model is correct. Similarly, the probability of observing the data under the alternate hypothesis is calculated by assuming a uniform distribution for the crack size. Then, the ratio

of these two likelihoods is calculated as the Bayes factor for each experimental data available for validation. The results of validation are tabulated in Table 5.5.

Table 5.5: Results of Model Validation

No. of Cycles	Crack size (mm)	Bayes Factor
10008	0.42	3.16
13852	0.31	2.06
14141	0.63	0.80
14368	0.41	2.35
14659	0.43	2.80
15155	1.54	0.70
15277	0.55	2.10
17746	0.43	2.56
17844	0.62	0.79
19912	0.35	3.00

From Table 5.5, it can be seen that the Bayes factor is greater than unity in most of the cases and lowest value is 0.70. The combined Bayes factor corresponding to all the data can be calculated using the method developed in Section 4.4.1 to be equal to 306. This corresponds to a 99.6% confidence in the model prediction.

5.7 Summary

This chapter applied the methods of Chapters III and IV for uncertainty quantification and model validation in fatigue crack growth analysis. Structures with complicated geometry and multi-axial variable amplitude loading conditions were considered. The finite element analysis used for calculation of stress intensity factor was replaced using a Gaussian process surrogate model. Different sources of uncertainty (physical variability, data uncertainty, and model error/uncertainty) were included in the crack growth analysis. Different types of model errors (discretization error, crack growth model error, surrogate model error) were considered explicitly. Deterministic

errors were corrected where they arose and stochastic errors were included by using random samples during uncertainty propagation.

Bayesian inference was used to calibrate the parameters of different models using inspection results (crack size after number of cycles, including cases where no crack was detected and crack was detected but size not measured). A Bayesian confidence metric was developed to assess the validity and quantify the confidence in fatigue crack growth prediction.

Note that this chapter used a particular set of models (modified Paris law, Wheelers retardation model, characteristic plane approach for the calculation of an equivalent stress intensity factor) only for the purpose of illustration. The crack growth analysis was limited to linear elastic fracture mechanics and planar cracks. Also, the effects of model assumptions such as planar crack, LEFM, equivalent stress intensity factor for multi-modal fracture, and crack retardation model were not individually quantified. However, the Bayesian framework for uncertainty quantification and model validation is general and can easily accommodate other advanced analysis models and corresponding model errors/uncertainty.

Though multiple models were used for crack growth prediction this chapter, they were not a part of a hierarchical system. In a hierarchical system, the goal is to to predict the system-level performance using subsystem-level, component-level, and system-level models. In this chapter, several models were used used to predict the crack growth in one component only. However, the method of Bayesian network is general enough to connect multiple models, irrespective of whether the models are present at the same level of hierarchy or at different levels of hierarchy. Hence, the Bayesian network provides a tool for integrating the various sources of uncertainty in hierarchical systems, which is the primary focus of the next chapter.

CHAPTER VI

INTEGRATION OF RESULTS FROM VERIFICATION, VALIDATION, AND CALIBRATION ACTIVITIES

6.1 Introduction

Chapter IV addressed the quantification of model uncertainty during various stages of the model development procedure. These various stages are grouped into five activities, as summarized below.

1. Development of conceptual model and the governing mathematical equations
2. Development of computer code to solve the above equations
3. Model verification - quantify solution approximation errors
4. Model calibration - quantify uncertainty in model parameters
5. Model validation - quantify extent of agreement between model prediction and experimental data

Chapter V demonstrated the various activities related to uncertainty quantification by computing solution approximation errors, model parameter uncertainty, and a validation metric equal to the probability that the model is supported by the experimental data. The results of calibration were used for model prediction, and this prediction was validated using the validation data. Suppose that the probability that the model is correct ($P(H_0|D)$ in Bayesian hypothesis testing and $P(M)$ in model reliability method) is equal to 0.6. How to make use of this probability for future prediction? Further, the model calibration was based on the model, and hence is inherently conditioned on the assumed model form. (If the KOH framework is pursued,

then the chosen model form corresponds to Eq. 4.9, and the result of calibration is conditioned on this model form.) The model validation procedure assigns a probability ($\frac{B}{B+1}$) for this model form to be correct.

Therefore, it can be seen that the methods proposed in Chapter IV address the three uncertainty quantification activities - model verification, validation, and calibration - only individually. It is necessary to integrate the results from these activities for the purpose of overall quantification in the model prediction. This is not trivial because of several reasons.

1. The solution approximation errors calculated as a result of the verification process need to be accounted for during calibration, validation, and prediction.

2. The result of validation may lead to a binary result, i.e. the model is accepted or rejected; however, even when the model is accepted, it is not completely correct, because the degree of correctness is indicated by the probability that the model is correct. Hence, it is necessary to account for this degree of correctness of the model in the prediction.

3. Model calibration is based on a particular model, and hence the posterior distributions of the calibrated parameters are inherently conditioned on the chosen model form. The challenge is to account for the probability that the model is correct in the uncertainty of the model parameters.

4. Third, calibration and validation are performed using independent data sets and it is not straightforward to compute their combined effect on the overall uncertainty in the response.

The issue of integrating information from various uncertainty quantification activities gets further complicated when system-level behavior is predicted based on a hierarchy of models. As the complexity of the system under study increases, there

may be several components and subsystems at multiple levels of hierarchy, which integrate to form the overall multi-level system. Each of these components and subsystems are represented using component-level and subsystem-level models which are mathematically connected to represent the overall system model which is used to study the underlying system. In each level, there is a computational model with inputs, parameters, and outputs, experimental data (hopefully available for calibration and validation separately), and several sources of uncertainty physical variability, data uncertainty (sparse or imprecise data, measurement errors), and model uncertainty (parameter uncertainty, solution approximation errors and model form error). In such a multi-level system, the first task would be to connect all the available models and associated sources of uncertainty.

Recent studies [145, 196, 197] by Mahadevan and co-workers at Vanderbilt University have demonstrated that the Bayesian network methodology provides an efficient and powerful tool to integrate multiple levels of models, associated sources of uncertainty and error, and available data at multiple levels. While the Bayesian approach can be used to perform calibration and validation individually for each model in the multi-level system, it is not straightforward to integrate the information from these activities in order to compute the overall uncertainty in the system-level prediction. This chapter extends the Bayesian approach to integrate and propagate information from verification, calibration, and validation activities in order to quantify the overall uncertainty in the system-level prediction and thereby directly aid in quantification of margins and uncertainties (QMU).

In Bayesian calibration, the goal is to estimate the probability distribution of the model parameter using the data available for calibration. Once the model is calibrated, it is validated using an independent set of input-output data; the Bayesian

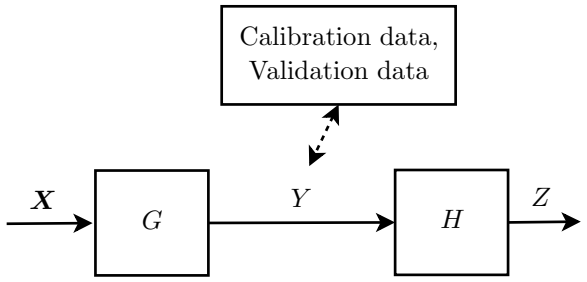
hypothesis testing method (Section 4.4.1) and the reliability-based method (Section 4.4.2) can be used for this purpose. Chapter IV discussed systematic methodologies for the inclusion of the different types of uncertainty in both calibration and validation activities; further, different types of data situations were also considered for both calibration and validation. While Bayesian inference and Bayesian hypothesis testing can be used for calibration (Section 4.3) and validation (Section 4.4) respectively, these two procedures are different and should not be confounded. The distinction will be clearly maintained in this dissertation; in fact, this chapter considers separate data sets for calibration and validation.

In Chapter IV, the methods for calibration and validation were demonstrated only for individual models with calibration and validation data. What happens when there is flow of information across multiple levels of models? Since the Bayesian approach represents all information through probability distributions, the problem reduces to propagating these probability distributions through the system hierarchy. The solution approximation errors, estimated as a result of model verification, can also be included as additional nodes in the Bayesian network. The resultant Bayesian network can be used for both the forward problem of uncertainty propagation [198] and inverse problem of calibration [199].[1] The results of calibration and validation activities are expressed in terms of PDFs of model parameters, and probability that each model is correct, respectively. The Bayesian method is thus able to integrate all the information from verification, calibration, and validation at multiple levels to calculate the overall system-level performance prediction uncertainty.

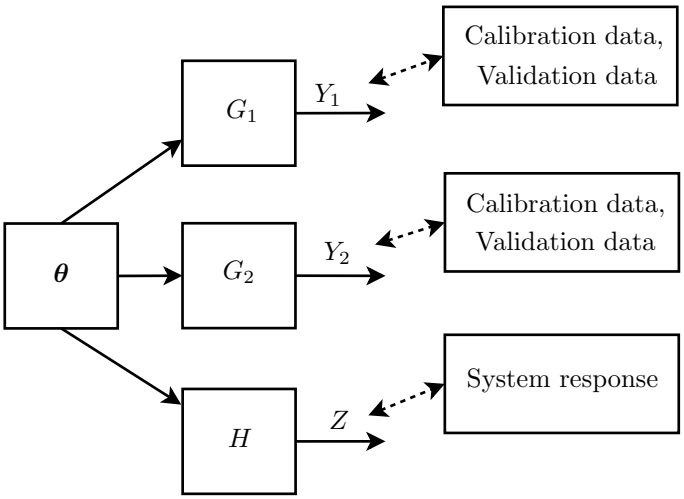
Recall that three different types of hierarchical configurations were discussed earlier in Section 1.3.1. The first two types of configurations – sequential (or feed-forward) and non-sequential – are considered in this chapter. In both cases, the

[1]Refer to [197–199] for details on the inclusion of all sources of uncertainty for both forward uncertainty propagation and parameter estimation.

quantity of interest is an overall system-level response, but there is a significant difference in how this quantity is calculated using information from uncertainty quantification activities (verification, validation, and calibration) performed at lower-levels using the corresponding models. The two types of configurations (sequential and non-sequential) are shown in Fig. 6.1.



(a) Sequential



(a) Non-Sequential

Figure 6.1: Two Types of Hierarchical Configurations

The third type of configuration discussed in Section 1.3.1, i.e. hierarchical system with feedback coupling, is not considered in this chapter. In such a system, two models (usually, governing two different physics) are interconnected in such a way that the input of one becomes an output to the other and vice versa. Chapter VII will consider

feedback coupling in multi-disciplinary systems, and propose an approach to replace feedback coupling with equivalent unidirectional coupling, thereby facilitating the use of Bayesian networks for systems with feedback coupling. The present chapter focuses on sequential and non-sequential system configurations, where the Bayesian network is used to integrate (1) the various sources of uncertainty across multiple levels of models; and (2) the results from uncertainty quantification activities (verification, validation, and calibration) in order to quantify the overall uncertainty in the system-level response.

When the hierarchy is "sequential", the output (Y) of a lower level model (G) becomes an input to a higher level model (H), as shown in Fig. 6.1 (a). Each model has its own set of model parameters (not indicated in Fig. 6.1 (a)); there may or may not be any model parameter common between two models. In this type of multi-level system, the uncertainty in the output of each model is propagated through higher levels of the hierarchy.

When the hierarchy is "non-sequential" as shown in Fig. 6.1 (b), the outputs (Y_1, Y_2, Z) and inputs across multiple models are not connected; rather, some model parameters ($\boldsymbol{\theta}$) are common to models at multiple levels of complexity. The model of the highest complexity (H) represents the system of interest and the system-level response (Z) needs to be calculated. The model parameters ($\boldsymbol{\theta}$) are calibrated using models and experiments of reduced complexity (e.g. isolated components or physics), and then propagated through the system model to compute the desired response. Urbina et al. [196] and Red-Horse and Paez [200] discuss such non-sequential models increasing complexity. Chapter VIII will also discuss non-sequential systems of increasing complexity and physics.

(Note: One could think of models of increasing complexity as sequential, but in this dissertation, the term "sequential" is used only to the case when the output of one model is input to the next model.)

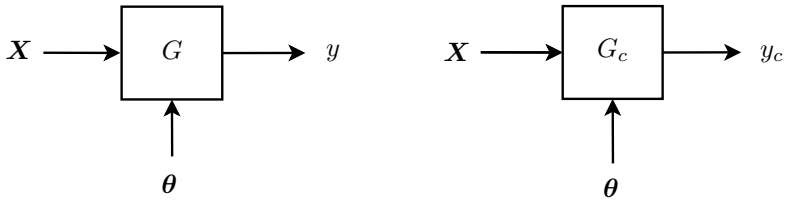
The proposed methods for integration of calibration, verification, and validation results are different for each type of configuration. In the sequential case, the linking variables are the outputs of the lower levels that become inputs to the higher level models, whereas in the non-sequential case, the linking variables are the common model parameters. With the focus on the linking variables, this chapter develops a Bayesian network-based methodology to integrate the results of verification, validation, and calibration activities, and compute the uncertainty in the overall system-level prediction.

The rest of this chapter is organized as follows. Section 6.2 discusses the proposed methodology for integration of calibration, verification, and validation for a single-level model. Sections 6.3 and 6.4 extend the proposed methodology for models with sequential and non-sequential configurations respectively. The proposed methods are illustrated using three numerical examples - a heat conduction example (single level model), an electric wire under heat conduction (two sequential models), and a structural dynamics problem (three non-sequential models) in Sections 6.5, 6.6 and 6.7 respectively.

6.2 Integration for a Single-level Model

Consider a single-level model as shown in Fig. 6.2. The inputs are $\boldsymbol{x}$, the model parameters are $\boldsymbol{\theta}$, the true solution of the mathematical equation is y, and the code output is y_c. Both y_c and y are deterministic functions of inputs ($\boldsymbol{x}$) and model parameters ($\boldsymbol{\theta}$).

This section proposes methods to integrate the results of calibration, verification,



(a) Mathematical Equation (b) Computer Code

Figure 6.2: A Single-level Model

and validation of the model. Since the process of verification is not related to data, it needs to be performed first; both calibration and validation must include the results of verification analysis (i.e. error quantification). Then, the results of verification, calibration, and validation are integrated to compute the overall uncertainty in the response quantity. Though the methods for verification, validation, and calibration were explained in detail in Chapter IV, they are briefly reviewed here. The reason is that verification, validation, and calibration were performed independently in Chapter IV, while now, they are performed in an integrated manner so as to quantify the uncertainty in the overall output.

6.2.1 Verification

The process of verification checks how close the code output is to the true solution of the mathematical equation. As stated earlier in Section 4.2, it is not only sufficient to verify that the two solutions are sufficiently close, but also essential to quantify the solution approximation error, i.e. the difference between the code output and true solution, in order to quantify the uncertainty in the prediction. As explained Sections 4.2.1 and 4.2.2, the solution approximation error is composed of both deterministic and stochastic terms. The discretization error arising in finite element analysis is deterministic, while the surrogate model error that arises as a result of replacing the finite element analysis with a surrogate model is stochastic. In the context

of uncertainty propagation, deterministic errors can be addressed by correcting the bias, and the corrected solutions are used to train the surrogate model; the stochastic errors of the surrogate model can be addressed through sampling based on their estimated distributions. As a result, the overall solution approximation error is also stochastic. The quantification of discretization error and surrogate model uncertainty was explained earlier in Section 4.2.

The true solution of mathematical equation can be computed as a function of the model inputs and parameters as $y(\boldsymbol{x};\boldsymbol{\theta}) = y_c(\boldsymbol{x};\boldsymbol{\theta}) + G_{se}(\boldsymbol{x};\boldsymbol{\theta})$. Since $G_{se}(\boldsymbol{x};\boldsymbol{\theta})$ is stochastic, y is stochastic even for given values of $\boldsymbol{x}$ and $\boldsymbol{\theta}$.

6.2.2 Calibration

The next step is to estimate the model parameters ($\boldsymbol{\theta}$) using input - output ($\boldsymbol{x}$ vs. y) data collected for calibration (D^C), using Bayes theorem as:

$$f_{\boldsymbol{\theta}}(\boldsymbol{\theta}|G, D^C) = \frac{L(\boldsymbol{\theta}) f_{\boldsymbol{\theta}}(\boldsymbol{\theta})}{\int L(\boldsymbol{\theta}) f_{\boldsymbol{\theta}}(\boldsymbol{\theta}) d\boldsymbol{\theta}} \tag{6.1}$$

In Eq. 6.1, $f_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ is the prior PDF and $f_{\boldsymbol{\theta}}(\boldsymbol{\theta}|G, D^C)$ is the posterior PDF; the calibration procedure uses the model form G and hence the posterior is conditioned on G. The function $L(\boldsymbol{\theta})$ is the likelihood of $\boldsymbol{\theta}$ defined as being proportional to the probability of observing the data D^C conditioned on the parameters $\boldsymbol{\theta}$. The calculation of this likelihood function was explained earlier in Section 4.3. Different cases of data were considered and the likelihood was for derived for each case in Section 4.3.8.

Note that the model "G" is used for calibration in Eq. 6.1; recall that "G" refers to the model corrected for discretization errors. Hence, the results of verification are included in the calibration procedure. Deterministic discretization errors are corrected for before training the surrogate model, and the surrogate model uncertainty

is included in the likelihood function as demonstrated by Kennedy and O'Hagan [201] and McFarland [52]. In addition, the construction of the likelihood function can also include additional uncertainty in other inputs and parameters, as explained in Section 4.3.8. Section 4.3.5 and Kennedy and O'Hagan [201] demonstrate how the model inadequacy function can be included in calibration and estimated along with the model parameters.

The posterior PDFs of the model parameters can be calculated using direct integration of the denominator in Eq. 6.1, if the number of calibration parameters is small. Alternatively, Markov Chain Monte Carlo sampling [202] methods such as Metropolis algorithm [38], Gibbs algorithm [203] or slice sampling [39] can be used to generate samples of the posterior distributions of the parameters. The method of slice sampling, explained earlier in Section 2.7, is used in this study.

6.2.3 Validation

Assume that an independent set of validation data (D^V) is available. The model is assumed to be verified and calibrated and the model prediction is compared against the validation data. The model prediction can be computed as a function of input as:

$$f_Y(y|\boldsymbol{x}, G, D^C) = \int f_Y(y|\boldsymbol{x}, \boldsymbol{\theta}) f_{\boldsymbol{\theta}}(\boldsymbol{\theta}|G, D^C) d\boldsymbol{\theta} \tag{6.2}$$

In the case of partially characterized validation data (e.g. field data), the input $\boldsymbol{x}$ may not be measured, in which case the model prediction must include the uncertainty in the input as:

$$f_Y(y|G, D^C) = \int f_Y(y|\boldsymbol{x}, \boldsymbol{\theta}) f_{\boldsymbol{X}}(\boldsymbol{x}) f_{\boldsymbol{\theta}}(\boldsymbol{\theta}|G, D^C) d\boldsymbol{\theta} \tag{6.3}$$

The above equations simply refer to uncertainty propagation through the model and hence the model prediction is conditioned on the event that the mathematical model is correct, and written as $f_Y(y|G, D^C)$. The results of verification are included while computing y; deterministic errors are addressed by correcting them and stochastic errors are addressed by sampling them. The results of calibration are included by using the posterior PDF of the model parameter in the prediction.

The model prediction is then compared with the validation data using Bayesian hypothesis testing; let $P(G)$ and $P(G')$ denote the probabilities that the model is correct (null hypothesis) and that the model is incorrect (alternate hypothesis) respectively. Prior to validation, if no information is available, $P(G) = P(G') = 0.5$. Using Bayesian hypothesis testing, these probabilities can be updated using the validation data (D^V), and the Bayes factor metric is defined as:

$$B = \frac{P(D^V|G)}{P(D^V|G')} \tag{6.4}$$

The likelihoods $P(D^V|G)$ and $P(D^V|G')$ are denoted as $L(G)$ and $L(G')$ respectively. The numerator $P(D^V|G)$ can be calculated using $f_Y(y|G)$ as:

$$L(G) \propto P(D^V|G) \propto \int f(D^V|y) f_Y(y|G, D^C) dy \tag{6.5}$$

In Eq. 6.5, the term $f(D^V|y)$ is calculated based on the error $\epsilon_m \sim N(0, \sigma^2)$. In order to compute $P(D^V|G')$, it is necessary to assume the alternate PDF $f_Y(y|G')$, i.e. the PDF of Y when the model is wrong. Expert opinion may be used to construct this PDF, or a uniform PDF may be used if no additional information is available. Then $L(G')$ is calculated similar to Eq. 6.5 by replacing $f_Y(y|G, D^C)$ with $f_Y(y|G')$. As explained earlier in Section 4.4.1, the probability that the model is correct, i.e. $P(G)$ can be calculated as $\frac{B}{B+1}$.

The Bayes factor is a probabilistic measure of model validity. Alternatively, $P(G|D^V)$ can also be calculated using the model reliability approach (Section 4.4.2) by choosing a suitable tolerance limit. This probabilistic measure for model validation facilitates propagation of the result of validation for system-level uncertainty quantification.

While the Bayesian hypothesis testing is one approach to calculate the probability that the model is correct, the reliability-based approach discussed earlier in Section 4.4.2 in Chapter IV provides an alternative methodology. Consider Eq. 4.48 where the observed data was denoted by z whereas now it is denoted by D^V. Similar to the Bayesian hypothesis testing method, the reliability-based method can also account for the different types of uncertainty, as explained earlier in Section 4.4.2, and the probability that the model is correct, i.e. $P(M)$ can be computed. The assumption of the alternate PDF $f_Y(y|G')$, though not necessary for the computation of $P(M)$, is still necessary for the purpose of integration of verification, validation, and calibration as explained in the following subsection.

6.2.4 Integration for Overall Uncertainty Quantification

The calibration procedure in Section 6.2.2 assumed that the model form G is correct and estimated the model parameters $\boldsymbol{\theta}$. In contrast, the validation procedure in Section 6.2.3 calculated the probability that the model G is correct by assuming the uncertainty in the model parameters $\boldsymbol{\theta}$. The two results can be combined to calculate the overall uncertainty in the model prediction, using the theorem of total probability as:

$$f_Y(y|D^C, D^V) = P(G|D^V) f_Y(y|G, D^C) + P(G') f_Y(y|G') \tag{6.6}$$

In Eq. 6.6, $P(G|D^V)$ may be replaced with $P(M)$ if the model reliability-based method is used instead of Bayesian hypothesis testing. Note that the result of verification, i.e. solution approximation error, was already included in both calibration and validation. Thus, the PDF $f_Y(y|D^C, D^V)$ includes the results of verification, calibration, and validation activities.

6.3 Sequential Configuration of Models

Consider a system which is studied using multiple levels of models and there is sequential information flow between these models, i.e., the output of a lower level model becomes an input to the higher level model, and hence is the linking variable between the two models. While the methods of verification, validation, and calibration can be applied to each of the individual models, the challenge is to integrate the results from these activities performed at multiple levels. This section proposes a methodology for the integration of verification, validation and calibration across multiple levels of modeling with sequential configuration. The proposed methodology is illustrated for two levels of models, as shown in Fig. 4, and Eq. 6.7, but can be extended to any number of levels of modeling without loss of generality.

$$\begin{aligned} Y &= G(\boldsymbol{X}; \boldsymbol{\theta}) \\ Z &= H(Y, \boldsymbol{W}; \boldsymbol{\alpha}) \end{aligned} \tag{6.7}$$

Assume that no data is available at the system level, i.e. it is not possible to validate/calibrate model H. Let D^C and D^V denote the data available on Y for calibration (of $\boldsymbol{\theta}$) and validation (of G) respectively. Let $\epsilon_m \sim N(0, \sigma_m^2)$ denote the measurement errors in the data.

The first step is to connect the various sources of uncertainty using a Bayesian network, as shown in Fig. 6.4.

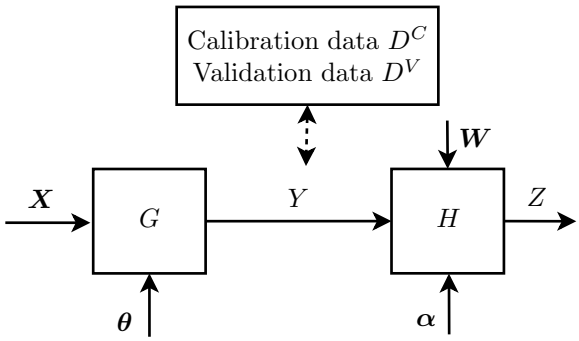


Figure 6.3: Sequential Information Flow: Two levels of Models

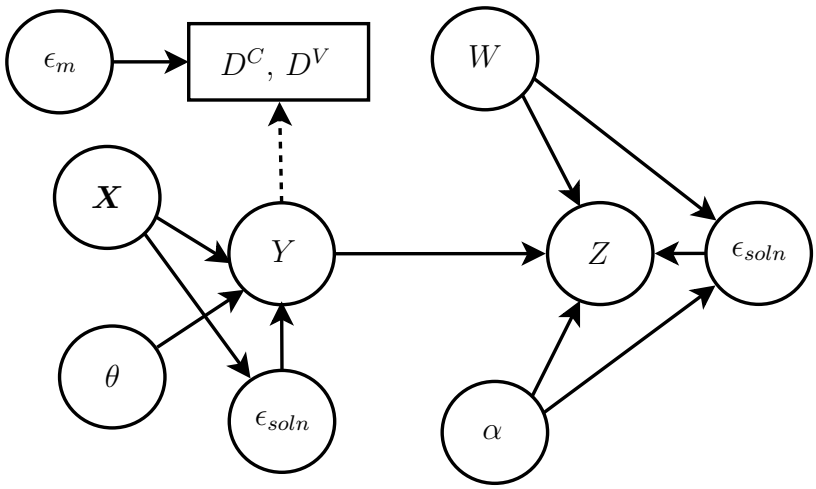


Figure 6.4: Bayesian Network: Sequential Configuration (Two Models)

This Bayesian network indicates that two sets of data are available for calibration and validation; the Bayesian methods for calibration and validation can be applied to these sets. If the KOH framework is pursued for calibration, then both the parameters and the model inadequacy function can be estimated. In addition to the parameters, the model inadequacy function would have also been included in this Bayesian network. However, it is still necessary to validate the overall model using additional validation data. In the validation stage, the Bayesian hypothesis approach or the model reliability method is pursued to quantify the probability that the model being correct.

In order to construct the Bayesian network, computer models are constructed to

solve Eq. 6.7. The task is to compute the overall uncertainty in Z by using lower level data; this uncertainty must include the effect of verification, calibration, and validation activities.

6.3.1 Verification, Calibration, and Validation

Both the models G and H can be verified since experimental data is not required for verification. During the process of verification, the solution approximation error (ϵ_{soln}) is quantified for both the models G and H. Note that the solution approximation error is a function of the inputs and the model parameters. As seen earlier in Section 6.2.1, the solution approximation error, in general, is a combination of deterministic and stochastic terms. The deterministic errors are addressed by correcting the FEA output, the corrected values are used to train the surrogate model; the surrogate model uncertainty is included as an additional node in the Bayesian network, as shown in Fig. 6.4. Note that these solution approximation errors (ϵ_{soln} for both G and H) account for the combined effect of both deterministic and stochastic errors. Now the Bayesian network includes quantification of solution approximation error and it can now be used for calibration, validation and system-level prediction.

The next step is to calibrate the model parameters. Suppose that the PDFs of the parameters $\boldsymbol{\theta}$ and $\boldsymbol{\alpha}$ are assumed to be $f_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ and $f_{\boldsymbol{\alpha}}(\boldsymbol{\alpha})$ before any testing; these are the prior PDFs. Since no data is available on Z, it is not possible to update the PDF of $\boldsymbol{\alpha}$. The data on Y, i.e. D^C is used to calibrate the parameters $\boldsymbol{\theta}$, using Bayesian inference, as in Section 6.2.2. The calibration procedure uses the data and assumes that the model is correct, and hence the posterior PDF of $\boldsymbol{\theta}$ is denoted by $f_{\boldsymbol{\theta}}(\boldsymbol{\theta}|G, D^C)$. During the calibration procedure, for every realization of $\boldsymbol{\theta}$, the corresponding solution approximation error is estimated and therefore, calibration is

based comparing y against experimental data, rather than y_c, thereby accounting for the results of verification during calibration.

Additional independent data (D^V) is assumed to be available for the purpose of validating the model G. The alternate hypothesis PDF $f_Y(y|G')$ is assumed and the posterior probability of model being correct, i.e. $P(G|D^V)$ is calculated as explained in Section 6.2.3; alternatively, the reliability-based metric $P(M)$ can also be used instead of $P(G|D^V)$.

6.3.2 Integration for Overall Uncertainty Quantification

The Bayesian network can be used for forward propagation of uncertainty using the principles of conditional probability and total probability. Prior to the collection of any data, the uncertainty in $\boldsymbol{x}$, $\boldsymbol{\theta}$, and $\boldsymbol{\alpha}$ can be propagated through the models as:

$$\begin{aligned} f_Z(z|H) &= \int f_Z(z|\boldsymbol{w},\boldsymbol{\alpha},y,H) f_{\boldsymbol{W}}(\boldsymbol{w}) f_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}) f_Y(y|G) d\boldsymbol{w} d\boldsymbol{\alpha} dy \\ f_Y(y|G) &= \int f_Y(y|\boldsymbol{x},\boldsymbol{\theta},G) f_{\boldsymbol{X}}(\boldsymbol{x}) f_{\boldsymbol{\theta}}(\boldsymbol{\theta}) d\boldsymbol{x} d\boldsymbol{\theta} \end{aligned} \tag{6.8}$$

However, this procedure assumes that (1) the PDFs of the parameters $\boldsymbol{\theta}$, and $\boldsymbol{\alpha}$ are $f_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ and $f_{\boldsymbol{\alpha}}(\boldsymbol{\alpha})$; and (2) the models G and H are correct. These two issues were addressed in calibration and validation respectively. While the PDF of $\boldsymbol{\alpha}$ did not change, the PDF of $f_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ was updated to $f_{\boldsymbol{\theta}}(\boldsymbol{\theta}|G,D^C)$. Further, the probability that G is correct, i.e. $P(G|D^V)$ was evaluated. These two quantities can now be used to calculate the overall uncertainty in Z. First, if the calibration data alone was used, then the PDFs of Y and Z are given by:

$$\begin{aligned} f_Z(z|G,H,D^C) &= \int f_Z(z|\boldsymbol{w},\boldsymbol{\alpha},y,H) f_{\boldsymbol{W}}(\boldsymbol{w}) f_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}) f_Y(y|G,D^C) d\boldsymbol{w} d\boldsymbol{\alpha} dy \\ f_Y(y|G,D^C) &= \int f_Y(y|\boldsymbol{x},\boldsymbol{\theta},G) f_{\boldsymbol{X}}(\boldsymbol{x}) f_{\boldsymbol{\theta}}(\boldsymbol{\theta}|G,D^C) d\boldsymbol{x} d\boldsymbol{\theta} \end{aligned} \tag{6.9}$$

The theorem of total probability can then be used to include the result of validation. The PDF of Y is modified as:

$$f_Y(y|D^C, D^V) = P(G|D^V) f_Y(y|G, D^C) + P(G'|D^V) f_Y(y|G') \tag{6.10}$$

The overall uncertainty in Z, which includes the results of verification, calibration, and validation, can be calculated as:

$$\begin{aligned} f_Z(z|H, D^C, D^V) &= P(G|D^V) f_Z(z|G, H, D^C) + P(G'|D^V) f_Z(z|G', H) \\ f_Z(z|G', H) &= \int f_Z(z|\boldsymbol{w}, \boldsymbol{\alpha}, y, H) f_{\boldsymbol{W}}(\boldsymbol{w}) f_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}) f_Y(y|G') d\boldsymbol{w} d\boldsymbol{\alpha} dy \end{aligned} \tag{6.11}$$

The PDF of Z is still conditioned on H because it is assumed that the model H is correct and it is not possible to calibrate/validate this model. In fact, Eq. 6.11 is equivalent to simply propagating the PDF $f_Y(y|D^C, D^V)$ (in Eq. 6.10) through the model H. Note that the model H has been verified; therefore, during uncertainty propagation, it is necessary to estimate and account for the solution approximation error, thereby including the result of verification of H. Thus, the PDF of the linking variable can be directly used to compute the uncertainty in the system-level response, thereby integrating the results of verification, validation, and calibration activities at a lower level.

6.3.3 Extension to Multiple Models

Until now, only the first model G was considered for verification, validation, and calibration. However, the proposed methodology is general and can be extended to multiple models. For example, consider the case where there are two models whose individual outputs become inputs for the system model. For example, consider the

equations:

$$\begin{aligned} Y_1 &= G_1(X_1, \theta_1) \\ Y_1 &= G_1(X_1, \theta_1) \\ Z &= H(Y_1, Y_2) \end{aligned} \tag{6.12}$$

The inputs to the models G_1 and G_2 are X_1 and X_2 respectively; the corresponding parameters are θ_1 and θ_2 respectively. The Bayesian network for this multi-level system is shown in Fig. 6.5.

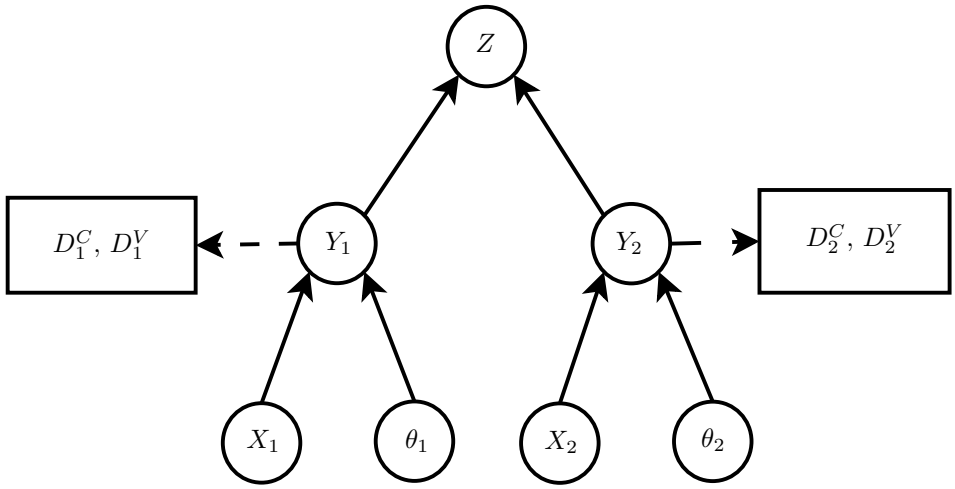


Figure 6.5: Bayesian Network: Sequential Configuration (Multiple Models)

Assume that there is no data at the system level Z, but data is available for calibration and validation of lower-level models G_1 and G_2, as shown in the Bayesian network in Fig. 6.5. Using the calibration data, the PDFs $f(\theta_1|G_1, D_1^C)$, $f(\theta_2|G_2, D_2^C)$, $f(y_1|G_1, D_1^C)$, and $f(y_2|G_2, D_2^C)$ are calculated. Using the validation data, the probability $P(G_1|D_1^V)$ and $P(G_2|D_2^V)$ are calculated; further $P(G_1'|D_1^V) = 1 - P(G_1|D_1^V)$ and $P(G_2'|D_2^V) = 1 - P(G_2|D_2^V)$. As explained earlier, the probabilities that G_1 and G_2 are correct can also be calculated using the reliability-based metric.

The unconditional PDF of Z needs to be calculated by considering four quantities:

1. $P(G_1 \cap G_2 | D_1^V, D_2^V) = P(G_1 | D_1^V) P(G_2 | D_2^V)$

2. $P(G_1 \cap G_2' | D_1^V, D_2^V) = P(G_1 | D_1^V) P(G_2' | D_2^V)$

3. $P(G_1' \cap G_2 | D_1^V, D_2^V) = P(G_1' | D_1^V) P(G_2 | D_2^V)$

4. $P(G_1' \cap G_2' | D_1^V, D_2^V) = P(G_1' | D_1^V) P(G_2' | D_2^V)$

Note the assumption that the two models G_1 and G_2 are independent. If the dependence is known, then it can be included in the calculation of the joint probabilities. Then, the unconditional PDF of Z is written as:

$$\begin{aligned} f_Z(z | D_1^C, D_1^V, D_2^C, D_2^V, H) \; &= P(G_1 | D_1^V) P(G_2 | D_2^V) f_Z(z | G_1, G_2, H) \\ &+ P(G_1' | D_1^V)) P(G_2 | D_2^V) f_Z(z | G_1', G_2, H) \\ &+ P(G_1 | D_1^V) P(G_2' | D_2^V) f_Z(z | G_1, G_2', H) \\ &+ P(G_1' | D_1^V) P(G_2' | D_2^V) f_Z(z | G_1', G_2', H) \end{aligned} \tag{6.13}$$

In Eq. 6.13, $f_Z(z|G_1, G_2, H)$ is calculated by propagating the posteriors of Y_1 and Y_2 through H, since both the models are correct; $f_Z(zG_1', G_2, H)$ is calculated by propagating the alternate PDF of Y_1 and the posterior of Y_2 through H, since only G_2 is correct; similarly, $f_Z(z|G_1, G_2', H)$ is calculated by propagating the posterior of Y_1 and alternate PDF of Y_2, and $f_Z(z|G_1', G_2', H)$ is calculated by propagating the alternate PDFs of Y_1 and Y_2. An alternative approach would be to simply compute the unconditional PDFs $f(y_1|D_1^C, D_1^V)$ and $f(y_2|D_2^C, D_2^V)$ similar to Eq. 6.10, and propagate these PDFs through the model H. Both the approaches will yield to the same resultant PDF of Z, which accounts for the results of verification, validation, and calibration activities in both the models G_1 and G_2.

Similar to the previous subsection (Section 6.3.2), the model H has been verified since verification does not need experimental data. While propagating the uncertainty through the model H, it is necessary to estimate and account for the solution

approximation error for every evaluation of H, thereby accounting for the result of verification of H in the procedure for system-level uncertainty quantification.

6.4 Non-Sequential Configuration of Models

Sometimes, a system model is developed using progressively complex models and corresponding experiments (isolated features, isolated physics, simplified geometry, scale models, etc.). The experiments of lowest complexity (simplest geometry or single physics) have been referred to as unit-level experiments [141]. A higher-level experiment could include an assembly of units or combined physics.

A typical example of such a system is discussed in [196], where material level tests (lowermost level), performance of a single joint, and performance of three joints are used to calibrate underlying material and model parameters that are used in the overall system-level model. Usually, in such a system, the complexity increases going up the hierarchy (more physics, features, components, etc.). Assume that the system-level model is given by:

$$Z = H(\boldsymbol{\theta}, \boldsymbol{X}, \boldsymbol{\Psi}) \tag{6.14}$$

In Eq. 6.14, Z is the system-level prediction, $\boldsymbol{\theta}$ is the set of model parameters which are calibrated based on lower level models and tests, $\boldsymbol{\Psi}$ is the set of additional model parameters at the system-level, and $\boldsymbol{X}$ are the inputs.

Consider two lower level models - first level G_1 and second level G_2. Both these models have common model parameters $\boldsymbol{\theta}$, but they have their own inputs ($\boldsymbol{X_1}$ and $\boldsymbol{X_2}$) and outputs (Y_1 and Y_2); in addition, they may have additional lower-level model

parameters ($\boldsymbol{\Psi_1}$ and $\boldsymbol{\Psi_2}$).

$$\begin{aligned} Y_1 &= G_1(\boldsymbol{\theta}, \boldsymbol{X_1}, \boldsymbol{\Psi_1}) \\ Y_2 &= G_2(\boldsymbol{\theta}, \boldsymbol{X_2}, \boldsymbol{\Psi_2}) \end{aligned} \tag{6.15}$$

Assume that separate sets of data are available for calibration (D_1^C and D_2^C for levels 1 and 2 respectively) and validation (D_1^V and D_2^V for levels 1 and 2 respectively). Full system testing is not possible, i.e. there is no test data is available at the system level (Z) and it is required to quantify the uncertainty in the system-level prediction using the data at the lower levels (Y_1 and Y_2). The inputs, model parameters, outputs, and data at all levels are connected through a Bayesian network, as shown in Fig. 6.6.

6.4.1 Verification, Calibration, and Validation

The steps of verification, calibration, and validation in each model are similar to the previous sections. It is possible to verify all the three models (G_1, G_2, and H) and compute the solution approximation error; deterministic errors are simply corrected where they occur and stochastic errors are included in the Bayesian network in Fig. 6.6.

If $\boldsymbol{\theta}$ is estimated using each individual model (G_1 or G_2) and the corresponding calibration data (D_1^C or D_2^C), then the corresponding PDFs of the model parameter $\boldsymbol{\theta}$ is $f(\boldsymbol{\theta}|D_1^C, G_1)$ or $f(\boldsymbol{\theta}|D_2^C, G_2)$ respectively. The Bayesian network facilitates the simultaneous use of both models and the corresponding data to calibrate $\boldsymbol{\theta}$ and obtain the PDF $f(\boldsymbol{\theta}|D_1^C, D_2^C, G_1, G_2)$. This step of simultaneous calibration using multiple data sets from experiments of differing complexity is different from the calibration considered in Sections 6.2 and 6.3, where only one model and the corresponding calibration data were used to estimate $\boldsymbol{\theta}$. In order to integrate the results of verification, validation, and calibration in Section 6.4.2 below, all the PDFs, i.e. those calibrated

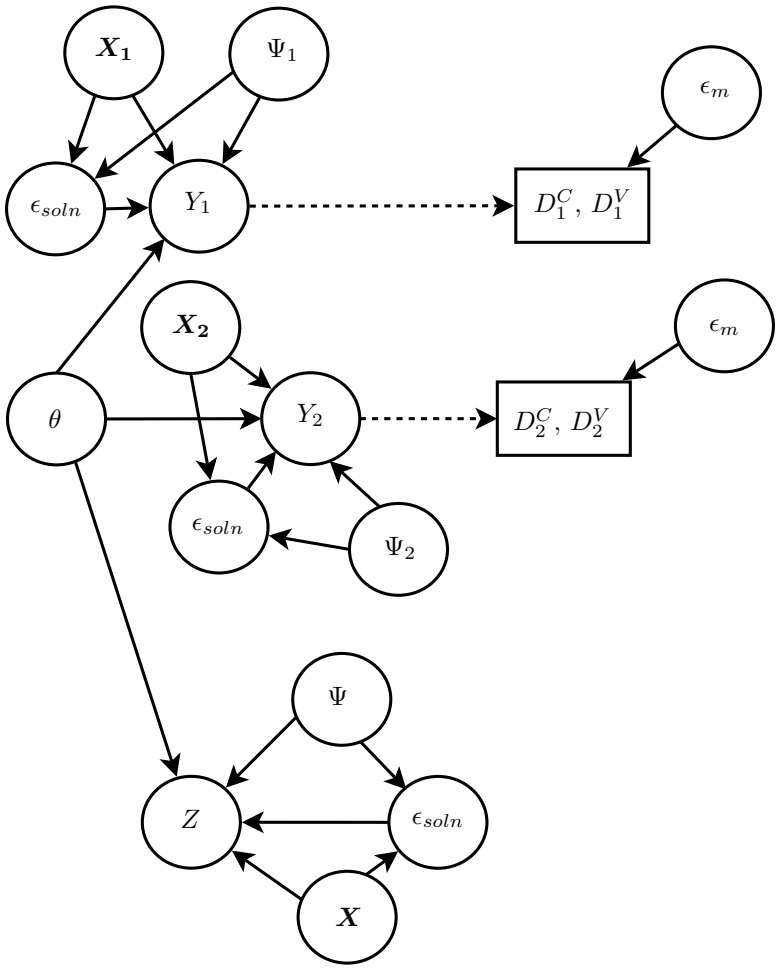


Figure 6.6: Bayesian Network: Non-sequential Configuration

using individual data sets ($f(\boldsymbol{\theta}|D_1^C, G_1)$ and $f(\boldsymbol{\theta}|D_2^C, G_2)$) as well as those calibrated using multiple data sets ($f(\boldsymbol{\theta}|D_1^C, D_2^C, G_1, G_2)$) are necessary.

The use of validation data is identical to the procedure in Sections 6.2 and 6.3. The quantities $P(G_1|D_1^V)$ and $P(G_2|D_2^V)$ are calculated using the Bayes factor metric; further $P(G_1'|D_1^V) = 1 - P(G_1|D_1^V)$ and $P(G_2'|D_2^V) = 1 - P(G_2|D_2^V)$. Alternatively, the reliability-based method can also be used to calculate this probability. Since the two models are assumed independent, $P(G_1 \cap G_2|D_1^V, D_2^V) = P(G_1|D_1^V)P(G_2|D_2^V)$.

6.4.2 Integration for Overall Uncertainty Quantification

The proposed method for overall uncertainty quantification and integration of the above activities is different from Section 6.3 because the linking variables in this case are the model parameters whereas the linking variables in Section 6.3 were the outputs of lower level models. While the unconditional PDF of the lower level output was calculated in Section 6.3, it is now necessary to calculate the unconditional PDF of the model parameter, that accounts for validation results. This is done using the total probability theorem as:

$$\begin{aligned} f_{\boldsymbol{\theta}}(\boldsymbol{\theta}|D_1^C, D_2^C, D_1^V, D_2^V) =& f_{\boldsymbol{\theta}}(\boldsymbol{\theta}|G_1 \cap G_2) P(G_1 \cap G_2|D_1^V, D_2^V) \\ &+ f_{\boldsymbol{\theta}}(\boldsymbol{\theta}|G_1' \cap G_2) P(G_1' \cap G_2|D_1^V, D_2^V) \\ &+ f_{\boldsymbol{\theta}}(\boldsymbol{\theta}|G_1 \cap G_2') P(G_1 \cap G_2'|D_1^V, D_2^V) \\ &+ f_{\boldsymbol{\theta}}(\boldsymbol{\theta}|G_1' \cap G_2') P(G_1' \cap G_2'|D_1^V, D_2^V) \end{aligned} \tag{6.16}$$

Eq. 6.16 is expressed as the sum of four terms; the first term $f_{\boldsymbol{\theta}}(\boldsymbol{\theta}|G_1 \cap G_2)$ is calculated for the case when both models are correct. Hence, both models can be used for calibration, and hence, $f_{\boldsymbol{\theta}}(\boldsymbol{\theta}|D_1^C, D_2^C, G_1, G_2)$ must be used. The second term $f_{\boldsymbol{\theta}}(\boldsymbol{\theta}|G_1' \cap G_2)$ means that the PDF of $\boldsymbol{\theta}$ is calculated for the case when model G_1 is wrong but model G_2 is correct. Hence, model G_1 should not be used for calibration, and hence, the resulting PDF is equal to $f_{\boldsymbol{\theta}}(\boldsymbol{\theta}|D_2^C, G_2)$. Similarly, the third term $f_{\boldsymbol{\theta}}(\boldsymbol{\theta}|G_1 \cap G_2')$ is equal to $f_{\boldsymbol{\theta}}(\boldsymbol{\theta}|D_1^C, G_1)$, and the fourth term $f_{\boldsymbol{\theta}}(\boldsymbol{\theta}|G_1' \cap G_2')$ is calculated under the condition that both G_1 and G_2 are wrong, and is simply equal to the prior $f_{\boldsymbol{\theta}}(\boldsymbol{\theta})$. The unconditional PDF $f_{\boldsymbol{\theta}}(\boldsymbol{\theta}|D_1^C, D_2^C, D_1^V, D_2^V)$ calculated in Eq. 6.16 accounts for the verification, calibration, and validation activities with respect to each of the lower level models. This unconditional PDF is propagated through the system model $H(\boldsymbol{\theta}, \boldsymbol{X}, \boldsymbol{\Psi})$, in order to quantify the uncertainty in the system-level response Z.

6.5 Example 1: Single-level Model

This section discusses a numerical example, where a single-level model is subject to verification, validation and calibration. The results of these activities are integrated to calculate the overall uncertainty in the response quantity.

6.5.1 Description of the Problem

Consider the steady state heat transfer in a thin wire of length L, with thermal conductivity k, and convective heat coefficient β. Assume that the heat source is $Q(x) = 25(2x - L)^2$, where x is measured along the length of the wire. For the sake of illustration, it is assumed that this problem is essentially one dimensional and that the solution can be obtained from the following boundary value problem [91].

$$\begin{aligned} -k\frac{\partial^2 T}{\partial T^2} + \beta T &= Q(x) \\ T(0) &= T_0 \\ T(L) &= T_L \end{aligned} \tag{6.17}$$

This problem was earlier considered for model validation under uncertainty in Section 4.4.4 in Chapter IV; in this chapter, the goal is perform integration of all uncertainty quantification activities to compute the overall uncertainty in the system response.

The length of the wire is assumed to be deterministic ($L = 4\ m$). The boundary conditions, i.e. the temperatures at the ends of the wire ($T(0)$ and $T(L)$) are assumed to be normally distributed with statistics N(0, 1). The thermal conductivity of the wire (k) is assumed to be normally distributed N(5, 0.2) with units $Wm^{-1}/^{\circ}C$.. The convective heat coefficient (β) is an unknown parameter which needs to be estimated using calibration data (D^C); this quantity is assumed to have a normally distributed

prior as N(0.5, 0.05). The goal of the model is to predict the temperature (Y) at the mid-point of the wire.

6.5.2 Verification, Validation, and Calibration

First, the differential equation in Eq. 6.17 is solved using a finite difference code. Three different discretization sizes are considered, and Richardson extrapolation [112] is used to calculate the solution approximation error which is used to correct the model prediction every time this differential equation is solved. Then, calibration data ($D^C = \{22; 23; 25; 26.1; 25.4\}$, in $^\circ C$) is assumed to be available and used to calibrate the unknown model parameter, i.e, the convective heat coefficient (β). During the calibration, the model inadequacy term was estimated to be insignificant and approximately equal to zero. The prior ($f_\theta(\theta)$) and posterior ($f_\theta(\theta|G, D^C)$) PDFs of β are shown in Fig. 6.7. Additional validation data ($D^V = \{24; 24.5; 24.6; 23.8\}$, in $^\circ C$) is used to compute the probability that the temperature prediction model is correct, i.e. $P(G) = 0.84$.

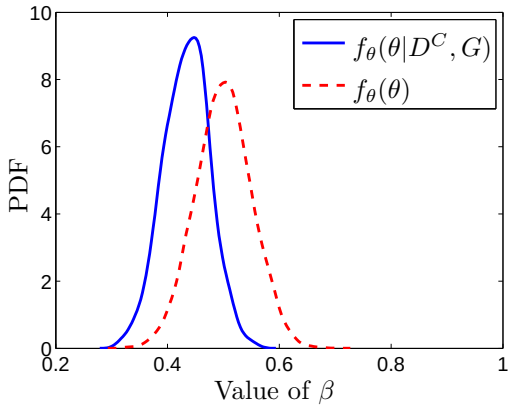


Figure 6.7: PDF of Convective Heat Coefficient (β)

6.5.3 Integration and Overall Uncertainty Quantification

The method developed in Section 6.2.4 is used to calculate the unconditional PDF of temperature using the principle of total probability, as shown in Fig. 6.8. This PDF integrates the results of verification, validation, and calibration to compute the overall uncertainty in the temperature at the mid-point of the wire.

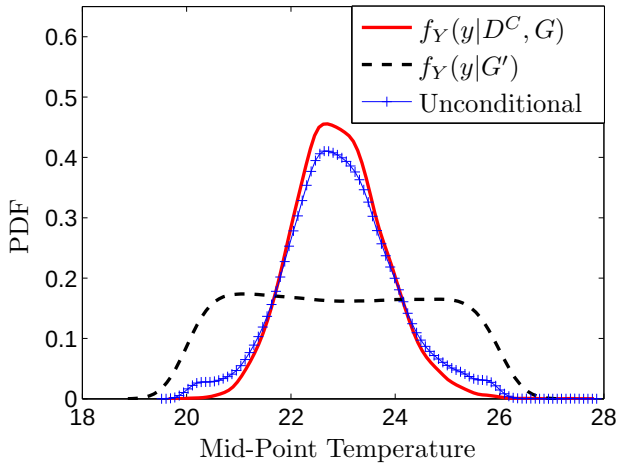


Figure 6.8: PDF of Mid-point Temperature

Fig. 6.8 indicates three PDFs (i) $f_Y(y|G, D^C)$ denotes the model prediction, (ii) $f_Y(y|G')$ denotes the prediction under the alternate hypothesis (assumed uniform; due to sampling errors and use of kernel density estimation for plotting, the PDF is not perfectly horizontal in Fig. 6.8), and (iii) $f_Y(y|G, D^C, D^V)$ which represents the PDF that integrates the validation result with the previous calibration and verification activities. The third PDF is referred to as the unconditional PDF of the temperature response, since it is not conditioned on the model form. Conventionally the model prediction alone is used for performance prediction and reliability analysis. The difference between the model prediction PDF and the unconditional PDF is prominent, especially in the tail region. For example, if the component is assumed to fail when the temperature is greater than 25 $°C$, then the model prediction PDF

gives the failure probability as 0.0135, whereas the unconditional PDF gives the failure probability as 0.0390. Thus, it is clear that the model prediction must not be used directly; it is necessary to include the results of verification, validation, and calibration in overall uncertainty quantification and performance prediction.

6.6 Example 2: Sequential Configuration

This section discusses a sequence of two models featuring thermal and electrical analyses. This example is an extension of the heat conduction problem in Section 6.5; the temperature rise in the wire causes change in the electrical resistance. The goal is to predict the system response, which is the electric current in the wire. Hence, the output of the lower level model (temperature predictor in Eq. 6.17), i.e. temperature, becomes an input to a higher level model (current predictor), thereby exhibiting sequential information flow, as shown in Fig. 6.9.

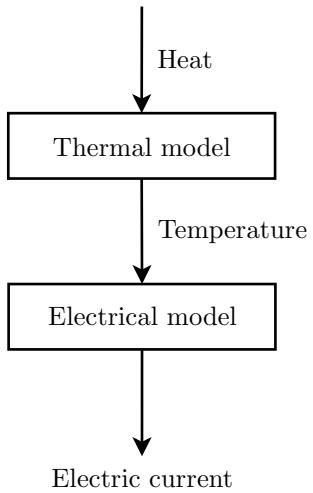


Figure 6.9: Thermal Electric Analysis

Consider the same wire as in Section 6.5. Before application of the heat, the resistance of the wire is given in terms of the resistivity (ρ), the cross section area

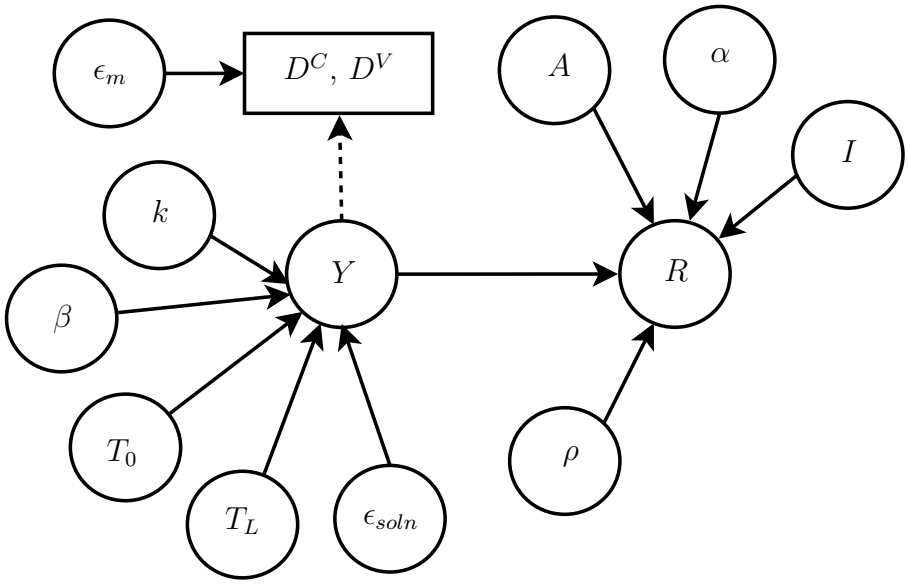


Figure 6.10: Bayesian Network: Thermal Electric Analysis

(A), and length (L) as:

$$R_{old} = \rho \frac{L}{A} \tag{6.18}$$

After steady state is reached, the mid-point temperature (Y) computed in Eq. 6.17 causes an increase in the resistance of the wire; this increase is evaluated using the coefficient of resistivity (α). The current through the wire when a 10V voltage is applied is calculated as:

$$I = \frac{10}{R_{old}(1 + \alpha Y)} \tag{6.19}$$

Assume that there is no electrical performance test data for the wire, and it is required to predict the uncertainty in the electrical current, by including the results of verification, validation, and calibration in the lower level model. The two models and the associated sources of uncertainty are connected through a Bayesian network as shown in Fig. 6.10.

Since the thermal model used for temperature prediction has already been verified, calibrated, and validated, the unconditional PDF of the temperature is simply

propagated through the current-predictor model to calculate the current in the wire. For the purpose of illustration, and to see the effect of uncertainty in Y on the uncertainty in electrical current (I), the other parameters of the current-predictor model (α, A, ρ) are chosen to be deterministic. The PDF of the current of the wire is shown in Fig. 6.11, for three cases.

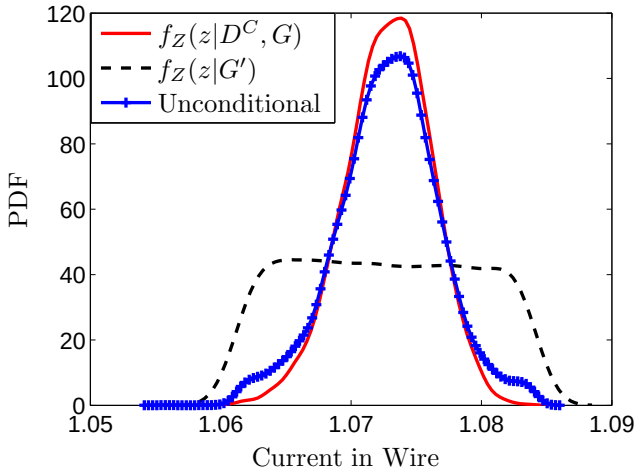


Figure 6.11: PDF of Current: System Response

The PDF $f_Z(z|G, D^C)$ is obtained by propagating the model prediction of thermal model through the electrical model, and the PDF $f_Z(z|G')$ is obtained by propagating the alternate PDF of temperature ($f_Y(y|G')$) through the electrical model. The unconditional PDF ($f_Z(z|D^C, D^V)$) represents the current response by integration of verification, validation, and calibration activities with respect to the lower level heat conduction model. Similar to the previous example, the difference between $f_Z(z|G, D^C)$ and the unconditional $f_Z(z|D^C, D^V)$ is prominent, especially in the tail region. For example, $1-F_Z(z = 1.08|G) = 0.0086$ whereas $1-F_Z(z = 1.08) = 0.0400$.

6.7 Example 3: Non-Sequential Configuration

This section illustrates the methodology for non-sequential hierarchical systems through a numerical example which consists of a three-level structural dynamics problem, as shown in Fig. 6.12. This numerical example was developed at Sandia National Laboratories [200], as a model validation challenge problem.

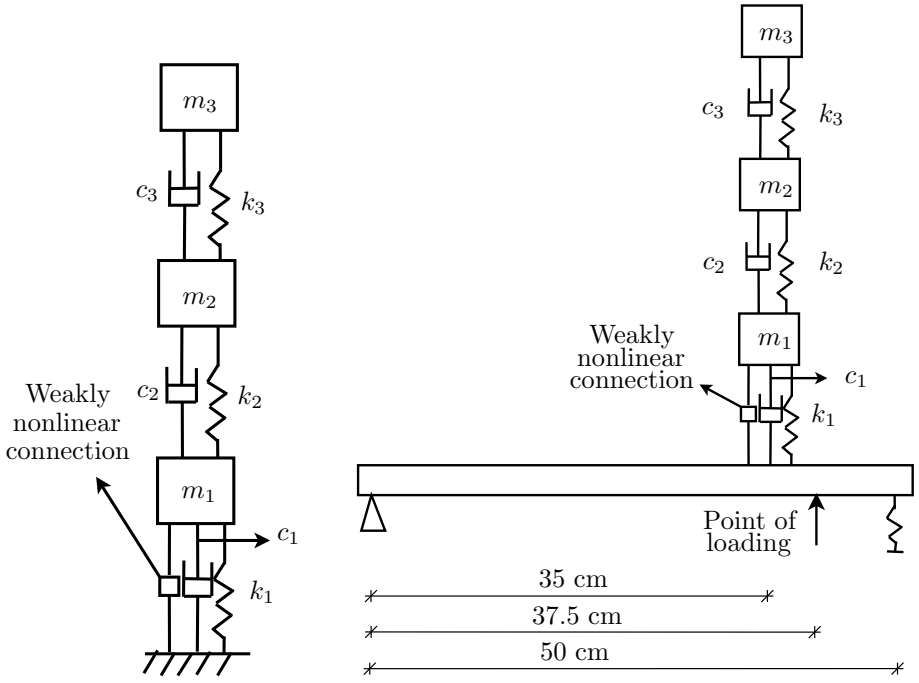


Figure 6.12: Multi-level Structural Dynamics Problem

6.7.1 Description of the Problem

In the first-level, three spring-mass-dampers are integrated to form a subsystem. In the second-level, the integrated spring-mass-damper subsystem is mounted on a beam to form the overall system. The overall objective is to compute the system-level output (R) which is defined to be the maximum acceleration of mass m_3, under a given realization of random process loading [200] on the beam. The model to

compute this system-level output is provided by Red-Horse and Paez [200]. This is the overall system-level model (denoted by Z); no test data is available at this level. The uncertainty in R needs to be computed based on information from lower level data. Two types of tests can be performed, at each of the levels.

1. **Level 1:** The three mass assembly in Fig. 6.12a is tested under sinusoidal loading (amplitude=10000 and angular velocity = 10 rad s^{-1}), and the acceleration of the top mass m_3 is measured. A model (denoted by G) is built to predict this response x_1. The construction of this model is straightforward and can be found in several textbooks [204]. Let D_1 denote test data; similar to the previous sections, two sets of test data are available: D_1^C for calibration, and D_1^V for validation.

2. **Level 2:** The beam with the 3-mass subsystem in Fig. 6.12b is tested under sinusoidal loading (amplitude=10000 and angular velocity = 10 rad s^{-1}), and the acceleration of the top mass m_3 is measured. A model (denoted by H) is provided in [200] to predict this response x_2. Similar to the previous sections, two sets of test data are available: D_2^C for calibration, and D_2^V for validation.

In this numerical example, for the sake of illustration, the stiffness values of the three masses, i.e. k_1, k_2, and k_3 are identified as the parameters to be calibrated using available test data. An additional set of data is used to validate the lower level models and all of this information is used to predict the system-level response R, defined earlier.

Prior distributions are assumed for k_1, k_2, and k_3 and later updated with test data to calculate posterior distributions. The system-level output R, in turn, is calculated by propagating the posterior distributions through the model Z. The numerical values (in SI units) of all the parameters are summarized in Table 6.1.

Table 6.1: Model Parameters: Structural Dynamics Problem

Number	Mass (m) (in kg)	Damping (c) (in Ns/m)	Prior Mean of Stiffness (μ_k) (in N/m)	Prior Std. Dev. of Mean (σ_k) (in N/m)
1	0.0125	0.023	5600	560
2	0.0193	0.021	11000	1100
3	0.0351	0.031	93000	9300

The mass of the beam is taken to be 0.1295 kg. Further numerical details of the beam are given in [200]. The model predictions, experimental data, errors, and the calibration quantities are connected using the Bayesian network, shown in Fig. 6.13.[1]

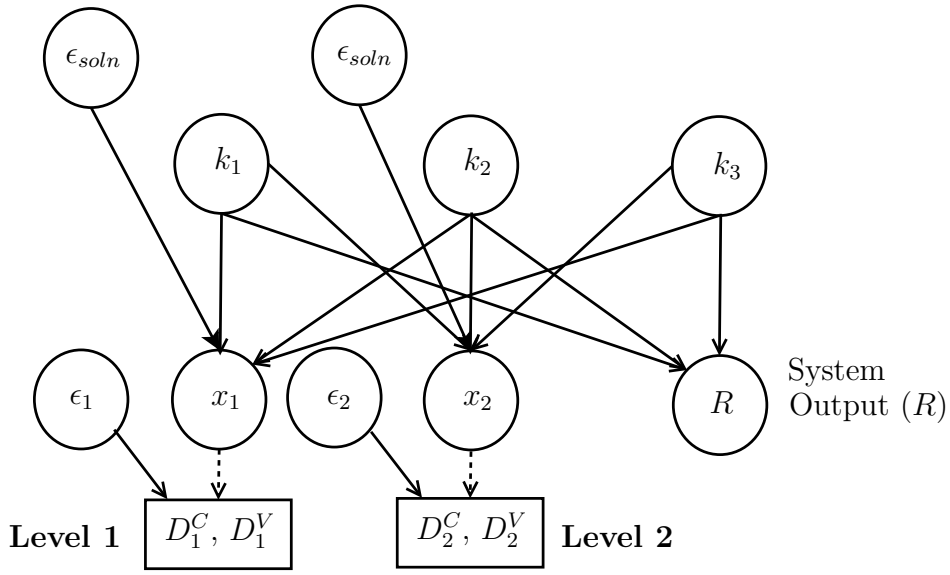


Figure 6.13: Bayesian Network: Structural Dynamics Problem

6.7.2 Verification, Calibration, and Validation

Two Gaussian process surrogate models are constructed to reduce the computational effort; the first is to replace the model H while the second is to compute the response quantity R. These surrogate models are constructed based on the computer codes provided in [200]. Whenever surrogate models are used for prediction,

[1]The same structural dynamics problem will be studied again in Chapter VIII, where three levels will be considered, and tests on individual models will also be included in the Bayesian network. See Fig. 8.5 in Section 8.4 in Chapter VIII.

there is a stochastic solution approximation error caused due to replacing the original model with a surrogate. This stochastic error is included in the Bayesian network in Fig. 6.13, through the nodes ϵ_{soln}. This error must be explicitly included in both the calibration and validation procedures.

The model parameters k_1, k_2 and k_3 are estimated using the calibration data, and shown in Figs. 6.14 – 6.16 respectively. The model inadequacy term was found to be insignificant and close to zero for this problem. All the four PDFs ($f_{\boldsymbol{\theta}}(\boldsymbol{\theta}|G \cap H)$, $f_{\boldsymbol{\theta}}(\boldsymbol{\theta}|G \cap H')$, $f_{\boldsymbol{\theta}}(\boldsymbol{\theta}|G' \cap H)$, and $f_{\boldsymbol{\theta}}(\boldsymbol{\theta}|G' \cap H')$ needed for the evaluation of Eq. 6.16 are also shown.

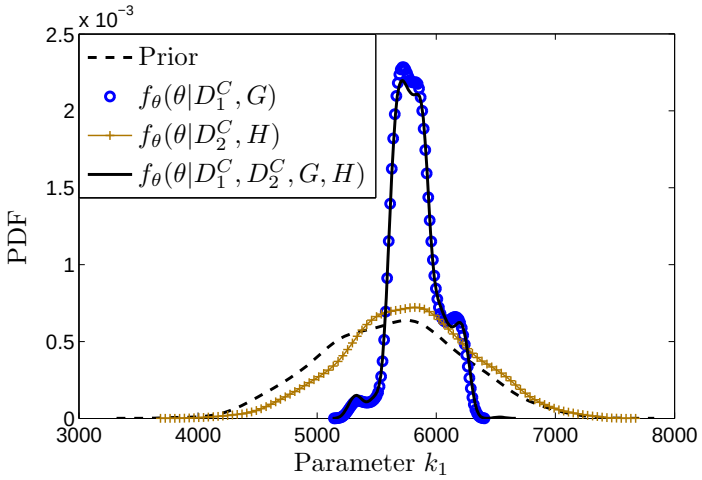


Figure 6.14: PDF of Parameter k_1

The next step is to validate the calibrated models. The models G and H are validated using two test measurements each, using the Bayesian hypothesis testing approach. The probabilities that the two models are correct are given by $P(G) = 0.25$ and $P(H) = 0.6$. It is assumed that the events that the models G and H are correct are independent; hence, $P(G \cap H) = 0.15$, $P(G \cap H') = 0.1$, $P(G' \cap H) = 0.45$, and $P(G' \cap H') = 0.3$. If the conditional probability that P(G is correct | H is correct) is available and not equal to P(G is correct), then this information can be included to calculate $P(G \cap H)$.

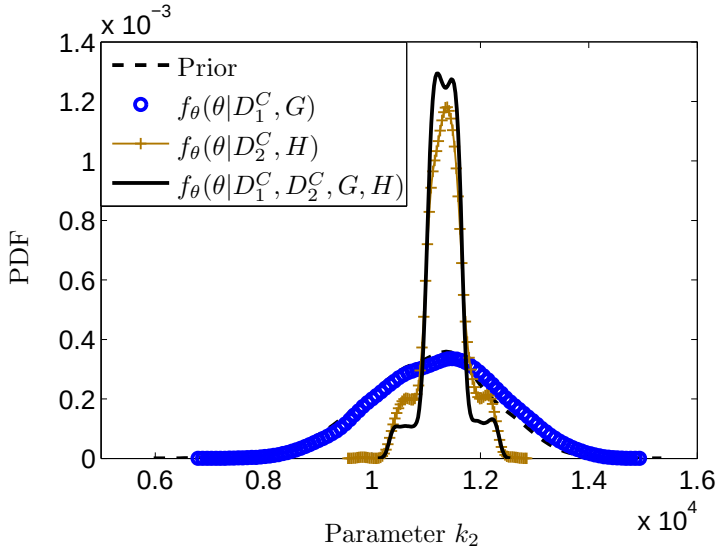


Figure 6.15: PDF of Parameter k_2

6.7.3 Integration for Overall Uncertainty Quantification

The next step is to calculate the unconditional PDFs of the calibration parameters k_1, k_2, and k_3, by including the validation result. This is accomplished by using Eq. 6.16, and the unconditional PDFs are used to compute the system-level response by propagating the uncertainty through the model Z; the resultant PDF is shown in Fig. 6.17.

Fig. 6.17 shows three PDFs; the first PDF ($f_R(r|G,H)$) is obtained by simply propagating the prior PDFs of the stiffnesses through models and hence is representative of all knowledge before test data collection. The second PDF ($f_R(r|D_1^C, D_2^C, G, H)$ includes the effect of verification (by considering surrogate model uncertainty) and calibration (by updating parameters using calibration data) but does not include the effect of validation (i.e. assumes the correctness of the lower-level models). The third PDF ($f_R(r|D_1^C, D_2^C, D_1^V, D_2^V)$ is the unconditional PDF and accounts for the results of verification, validation, and calibration activities in the lower level models. Similar

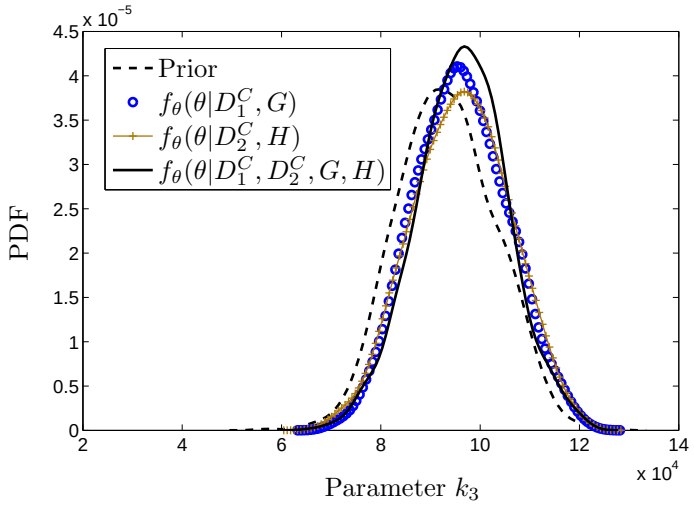


Figure 6.16: PDF of Parameter k_3

to the previous section, the differences between the PDFs are prominent in the tail region, and hence have a larger impact on reliability calculations.

6.8 Summary

Verification, validation, and calibration are significant activities in the process of model development. While methods for individual activities ere addressed in Chapter IV, the quantification of the combined effect of these activities on the overall system-level prediction uncertainty is addressed in this chapter.

Two independent sets of test data are considered: the first set is used to calibrate the model parameters and the second set is used to validate the calibrated model. The method of Bayesian inference developed in Section 4.3 is used to for calibration of model parameters. In order to validate the model, either the Bayesian hypothesis testing developed in Section 4.4.1 or the reliability-based method developed in Section 4.4.2 can be used. Both of these methods quantify the probability that the model is correct; recall that the reliability metric is more absolute than the Bayes factor metric.

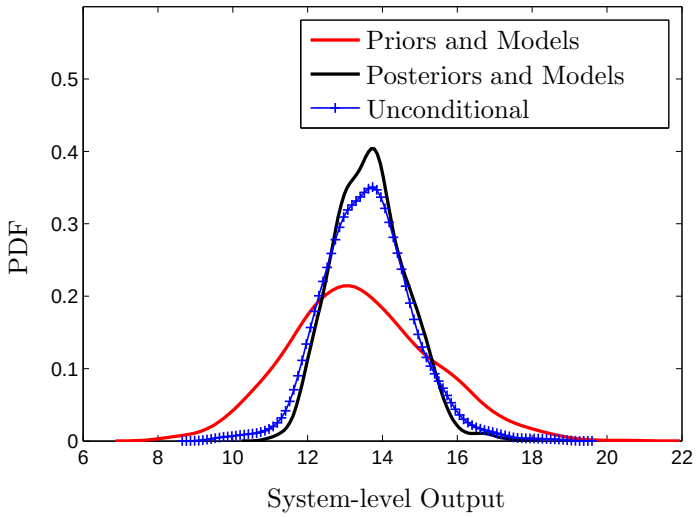


Figure 6.17: PDF of System Output R

This chapter proposed a methodology to integrate the results from verification, calibration, and validation in order to quantify the overall uncertainty in the system-level prediction. The integration methodology is then extended to two types of configurations – (1) multiple models with sequential configuration where the output of a lower-level model becomes an input to the higher-level model; and (2) multiple models with non-sequential configuration, where models of increasing complexity are considered and the system-level model parameters are inferred using test data at lower levels. The procedure for roll-up of calibration and validation results at lower levels is different for the two types of configurations; in the sequential system, the key idea is to compute the unconditional PDF of the output of the lower-level system, whereas in the non-sequential case, the key idea is to compute the unconditional PDF of the underlying model parameters. If a system-level prediction is based on models with both types of configurations (sequential and non-sequential), then the unconditional PDFs of the intermediate output and the parameters can both be used to compute the uncertainty in the overall system-level prediction uncertainty.

The proposed methodology offers considerable promise towards the quantification

of margins and uncertainties in multi-level system prediction. While calibration and validation have previously been performed independently at individual levels, this methodology systematically integrates all such activities in order to compute the system-level prediction uncertainty, thereby aiding in risk-informed decision making using all available information.

Several practical systems are multi-disciplinary, and there may be feedback coupling between the subsystem models, whereas the Bayesian network only allows acyclic dependencies between variables. The following chapter develops a new likelihood-based methodology for uncertainty quantification in such multi-disciplinary systems with feedback coupling. As a result of this new methodology, the feedback coupling can be replaced with a uni-directional coupling, thereby rendering the system configuration "sequential". Therefore, it will be straightforward to extend the methods developed in this chapter to systems with feedback coupling.

CHAPTER VII

UNCERTAINTY QUANTIFICATION IN MULTI-DISCIPLINARY SYSTEMS

7.1 Introduction

Practical engineering systems are often composed of multiple models and different types of physics interaction. In order to quantify the uncertainty in the system-level response, it is necessary to propagate the different sources of uncertainty through these multi-disciplinary models. As explained in Chapter I, these models can interact/combine in three different ways, leading to three types of system configurations, namely non-sequential, sequential and feedback configurations. Chapter VI addressed uncertainty quantification in non-sequential and sequential configurations through the use of a Bayesian network. However, the Bayesian network is acyclic and does not explicitly permit feedback coupling. This chapter proposes a new methodology for uncertainty quantification in multi-disciplinary systems with feedback coupling, which eventually facilitates the application of Bayesian networks to such configurations.

Multi-disciplinary systems analysis and optimization is an extensive area of research, and numerous studies in the literature have dealt with the various aspects of coupled multi-disciplinary analysis in several engineering disciplines. Researchers have focused both on the development of computational methods [205, 206] and the application of these methods to several types of multi-physics interaction, for example, fluid-structure [207], thermal-structural [208], fluid-thermal-structural [209], etc. Studies have considered these methods and applications either for multi-disciplinary analysis (MDA) or for multi-disciplinary optimization (MDO).

Computational methods for MDA can be classified into three different groups of

approaches [210]. The first approach, known as the field elimination method [210], eliminates one or more coupling variables (referred to as "field" in the literature pertaining to fluid-structure interaction) using reduction/elimination techniques such as integral transforms, model reduction, etc. This approach is restricted to linear problems that permit efficient and evident coupling. The second approach, known as the monolithic method [210, 211], solves the coupled analysis simultaneously using a single solver (for e.g. Newton-Raphson). The third approach, known as the partitioned method, solves the individual analyses separately with different solvers. The well-known fixed point iteration approach (repeated analysis until convergence of coupling variables), and the staggered solution approach [210, 212] are examples of partitioned methods. While the field elimination and monolithic methods tightly couple the multi-disciplinary analyses together, the partitioned method does not.

Two major types of methods have been pursued for MDO single level approaches and multi-level approaches. Single level approaches [205] include the multi-disciplinary feasible (MDF) approach (also called fully integrated optimization or the all-in-one approach), the all-at-once (AAO) approach (also called simultaneous analysis and design (SAND)), and the individual disciplinary feasible (IDF) approach. Multi-level approaches for MDO include collaborative optimization [213, 214], concurrent subspace optimization [215, 216], bi-level integrated system synthesis [217], analytical target cascading [218, 219], etc.

An important factor in the analysis and design of multi-disciplinary systems is the presence of uncertainty in the system inputs. It is necessary to account for the various sources of uncertainty in both MDA and MDO problems. The MDA problem focuses on uncertainty propagation to calculate the uncertainty in the outputs. In the MDO problem, the objective function and/or constraints may become stochastic

if the inputs are random. The focus of the present chapter is only on uncertainty propagation in multi-disciplinary analysis and not on optimization.

While most of the aforementioned methods for deterministic MDA can easily be extended to non-deterministic MDA using Monte Carlo sampling, this may be computationally expensive due to repeated evaluations of disciplinary analyses. Hence, researchers have focused on developing more efficient alternatives. Gu et al. [220] proposed worst case uncertainty propagation using derivative-based sensitivities. Kokkolaras et al. [218] used the advanced mean value method for uncertainty propagation and reliability analysis, and this was extended by Liu et al. [219] by using moment-matching and considering the first two moments. Several studies have focused on uncertainty propagation in the context of reliability analysis. Du and Chen [221] included the disciplinary constraints in the most probable point (MPP) estimation for reliability analysis. Mahadevan and Smith [222] developed a multi-constraint first-order reliability method (FORM) for MPP estimation. While all the aforementioned techniques are probabilistic, non-probabilistic techniques based on fuzzy methods [223], evidence theory [65], interval analysis [224], etc. have also been studied for MDA under uncertainty.

Similar to MDA, methods for MDO under uncertainty have also been investigated by several researchers. Kokkolaras et al. [218] extended the analytical target cascading approach to include uncertainty. A sequential optimization and reliability analysis (SORA) framework was developed by Du et al. [225] by decoupling the optimization and reliability analyses. Chiralaksanakul and Mahadevan [226] integrated solution methods for reliability-based design optimization with solution methods for deterministic MDO problems to address MDO under uncertainty. Smith [227] combined the techniques in [222] and [226] for the design of aerospace structures. The

literature on MDO under uncertainty is large; however, since the focus of this chapter is on MDA under uncertainty, aspects of MDO will not be considered; the rest of the chapter will focus on methods for uncertainty propagation in multi-disciplinary analysis.

Review of the above studies reveals that the existing methods for MDA under uncertainty are either computationally expensive or based on several approximations. Computationally expense is incurred in the following ways:

1. Using deterministic MDA methods with Monte Carlo sampling [27] require several thousands of evaluations of the individual disciplinary analyses.

2. Non-probabilistic techniques [65, 223, 224] use interval-analysis based approaches, which also require substantial computational effort. Further they are also difficult to interpret in the context of reliability analysis; this is an important consideration for MDO which may involve reliability constraints.

Approximations are introduced in the following manner:

1. Probability distributions are approximated with the first two moments [218, 219, 221, 222].

2. Approximations of individual disciplinary analyses may be considered using derivative-based sensitivities [220] or linearization at MPP for reliability calculation [221, 222].

Some of these problems can be overcome by the use of a decoupled approach that has been advocated by Du and Chen [221] and Mahadevan and Smith [222]. In this decoupled approach, Taylors series approximation and the first-order second moment (FOSM) method have been proposed to calculate the probability density function (PDF) of the coupling variables.

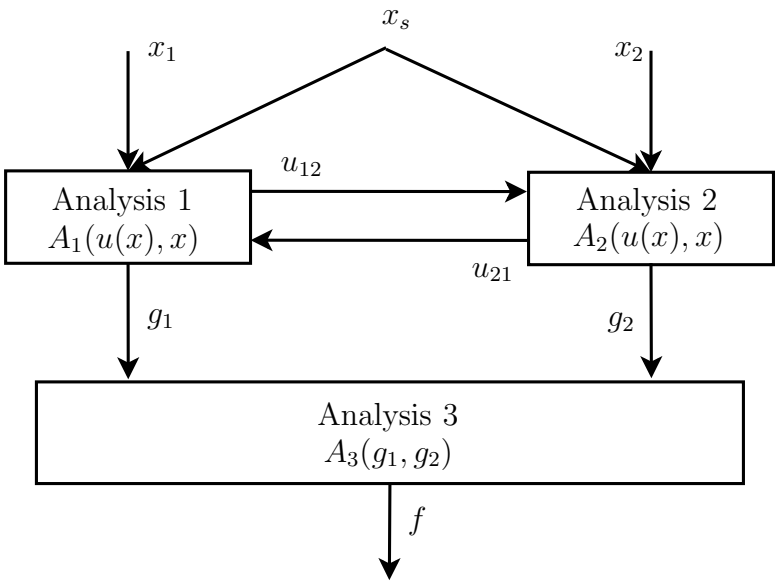


Figure 7.1: A Multi-disciplinary System

For example, consider the multi-disciplinary system shown in Fig. 7.1. Here $\boldsymbol{x} = \{x_1, x_2, x_s\}$ are the inputs, and $u(x) = \{u_{12}, u_{21}\}$ are the coupling variables. Note that this is a not only a multi-disciplinary system, but also a multi-level system where the outputs of the coupled analysis (g_1 and g_2) are used to compute a higher level system output (f).

Once the PDFs of the coupling variables u_{12} and u_{21} are estimated using the decoupled approach, the coupling between "Analysis 1" and "Analysis 2" is removed. In other words, the variable u_{21} becomes an input to "Analysis 1" and the variable u_{12} becomes an input to "Analysis 2", and the dependence between the quantities u_{12}, u_{21} and $\boldsymbol{x}$ is not considered any further. This "fully decoupled" approach reduces the computational effort considerably by avoiding repeated evaluations of the fully coupled system; however, this is still based on approximations and more importantly, suitable only when the aim is to estimate the statistics of g_1 or g_2.

In the case of a multi-level system, where the multi-disciplinary outputs (g_1 and g_2 in this case) could be inputs to another model (Analysis 3 in Fig. 7.1), the fully

decoupled approach will not be applicable for the following reason. In Fig. 7.1, for a given x, there is a unique g_1, and a unique g_2; in addition, for a given u_{12}, there is a unique u_{21}, and hence for a given g_1, there is a unique g_2. This functional dependence between u_{12} and u_{21}, and hence between g_1 and g_2, cannot be ignored when estimating the probability distribution of f. In the fully decoupled approach, the functional dependence between u_{12} and u_{21} is not preserved in subsequent analysis; once the PDFs of u_{12} and u_{21} are estimated, independent samples of u_{12} and u_{21} are used to generate samples of g_1 (using only Analysis 1) and g_2 (using only Analysis 2) which in turn are used to compute the statistics of f. This will lead to an erroneous estimate of f, since g_1 and g_2 values are not related to each other as they should be in the original system. This "subsequent analysis" need not necessarily refer to a higher level output; this could even refer to an optimization objective which is computed based on the values of g_1 and g_2 (or even u_{12} and u_{21}). Thus, if the objective is only to get the statistics of g_1 and g_2 as considered in [221, 222], then the fully decoupled approach is adequate. But if g_1 and g_2 are to be used in further analysis, then the one-to-one correspondence between u_{12} and u_{21} (and hence between g_1 and g_2) cannot be maintained in the fully decoupled approach. Hence, one would have to revert to the expensive Monte Carlo simulation outside a deterministic MDA procedure to compute the statistics of the output f. Thus, it becomes essential to look for alternatives to the fully decoupled approach, especially when the complexity of the system increases.

This chapter proposes a new likelihood-based approach for uncertainty propagation analysis in multi-level, multi-disciplinary systems. In this method, the probability of satisfying the inter-disciplinary compatibility is calculated using the principle of likelihood, which is then used to quantify the probability density function (PDF) of the coupling variables. The proposed approach offers several advantages:

1. The proposed method for the calculation of the PDF of the coupling variable

is theoretically exact; the uncertainty in the inputs is accurately propagated through the disciplinary analyses in order to calculate the PDF of the coupling variable. No approximations of the individual disciplinary analyses or the moments of the coupling variable are necessary.

2. This approach requires no coupled system analysis, i.e. repeated iteration between individual disciplinary analyses until convergence (as in fixed point iteration), thereby improving the computational cost.

3. For multi-level systems, the difficulty in propagating the uncertainty in the feedback variables to the system output is overcome by replacing the feedback coupling with unidirectional coupling, thereby preserving the functional dependence between the individual disciplinary models. The direction of coupling can be chosen either way, without loss of generality. This semi-coupled approach is also useful in an optimization problem where the objective function is a function of the disciplinary outputs.

The fact that the bi-directional coupling can be replaced with uni-directional coupling reduces the feedback coupling to feed-forward coupling, thereby rendering the system with a sequential configuration. Therefore, the methods for calibration, verification, validation and integration developed in Chapter VI can easily be extended to multi-disciplinary systems too. Since the method for integration of results from uncertainty quantification activities was developed earlier in Section VI, this chapter focuses only on the new likelihood-based methodology for uncertainty propagation, and explains how the bi-directional coupling can be replaced with uni-directional coupling.

The following sections describe the proposed likelihood-based methodology in detail. Section 7.2 discusses a "sampling with optimization-based deterministic MDA"

approach, which is an example of using the partitioned method along with Monte Carlo simulation. The proposed likelihood approach for MDA (LAMDA) is developed in Section 7.3 and its numerical implementation is discussed in Section 7.4. Section 7.5 illustrates the proposed methodology using a mathematical example and Section 7.6 uses the proposed methodology for a three-discipline analysis of a fire detection satellite [228].

7.2 Sampling outside Deterministic MDA

Consider the multi-disciplinary system shown earlier in Fig. 7.1. The overall goal is to estimate the probability distribution of the outputs g_1, g_2, and f, given the probability distributions of the inputs $\boldsymbol{x}$. As explained in Section 7.1, an intermediate step is to calculate the PDFs of the coupling variables u_{12} and u_{21} and then use these PDFs for uncertainty propagation.

First consider the deterministic problem of estimating the converged u_{12} and u_{21} values corresponding to given values of $\boldsymbol{x}$. The conventional fixed point iteration approach starts with an arbitrary value of u_{12} as input to "Analysis 2" and the resultant value of u_{21} serves as input to "Analysis 1". If the next output from "Analysis 1" is the same as the original u_{12}, then the analysis is said to have reached convergence and the inter-disciplinary compatibility is satisfied. However, if it is not, the conventional fixed point iteration approach treats the output of "Analysis 1" as input to "Analysis 2" and the procedure is repeated until convergence.

This search for the convergent values of u_{12} and u_{21} can be performed in an intelligent manner by formulating it as an optimization problem. For this purpose, define a new function G whose input is the coupling variable u_{12}, in addition to $\boldsymbol{x}$. The output of "G" is denoted by U_{12}, which is obtained by propagating the input through "Analysis 2" followed by "Analysis 1", as shown in Fig. 7.2.

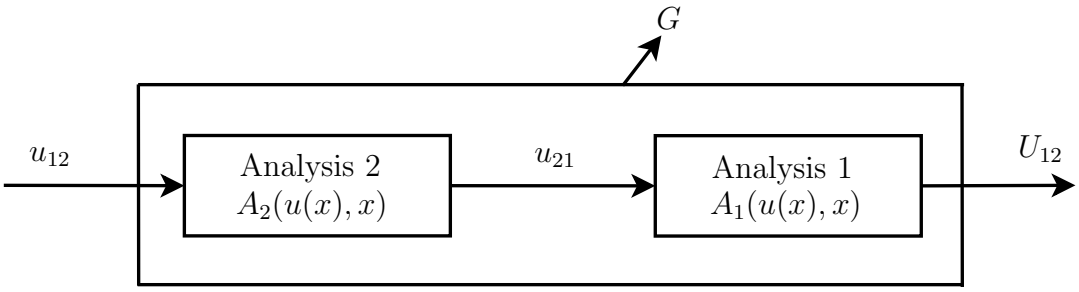


Figure 7.2: Definition of G

The multi-disciplinary constraint is said to be satisfied if and only if $u_{12} = U_{12}$. For a given $\boldsymbol{x}$, the convergent value of the coupling variable u_{12} can be obtained by minimizing the squared error $E = (u_{12} - G(u_{12}, \boldsymbol{x}))^2$ for a given set of inputs $\boldsymbol{x}$, where G is given by:

$$U_{12} = G(u_{12}, \boldsymbol{x}) = A_1(u_{21}, \boldsymbol{x}) \text{ where } u_{21} = A_2(u_{12}, \boldsymbol{x}) \tag{7.1}$$

Note that this is an unconstrained optimization problem. If the multi-disciplinary compatibility is satisfied, then $u_{12} = U_{12}$, and the optimum value of E will be equal to zero. In the rest of this chapter, it is assumed that it is possible to satisfy inter-disciplinary compatibility for each realization of the input $\boldsymbol{x}$; in other words, the multi-disciplinary analysis has a feasible solution for each input realization. Once the converged value of u_{12} is estimated, then the bi-directional coupling can be removed and replaced with a uni-directional coupling from "Analysis 2" to "Analysis 1" as shown in Fig. 7.3.

If there are multiple coupling variables in one direction, i.e. if u_{12} is a vector instead of a scalar, then E is also a vector, i.e. $E = [E_1, E_2, E_3, \ ... \ E_n]$. If the multi-disciplinary analysis has a solution, then the optimal value of the vector u_{12} will lead to $E_i = 0$ for all i's. Since each $E_i = 0$ by definition, the optimal value of

u_{12} can be estimated by minimizing the sum of all E_i's (instead of minimizing each E_i separately), and the minimum value of this sum will also be equal to zero.

This is a minor modification to the fixed point iteration approach; here the convergent value of the coupling variable is calculated based on an optimization which may choose iterations judiciously in comparison with the fixed point iteration approach. Hence, in terms of uncertainty propagation, the computational cost is still very high. The input values need to be sampled and for each realization, this optimization needs to be repeated and the entire distribution of the coupling variable needs to be calculated using many such samples.

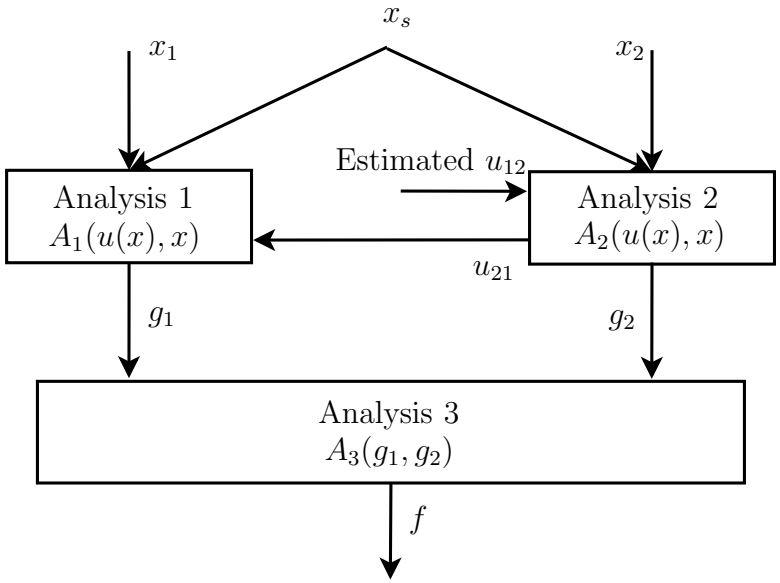


Figure 7.3: Partially Decoupled Multi-disciplinary System

Hereon, this approach is referred to as "sampling with optimization-based deterministic MDA" (SOMDA). Since this approach is still computationally expensive, a likelihood-based approach for MDA is proposed in the next section This approach does not require sampling and provides an efficient and theoretically accurate method for uncertainty propagation in multi-disciplinary analysis.

7.3 Likelihood-based Approach for MDA

The optimization discussed in the previous section is much similar to a least-squares based optimization; the difference being that a typical least squares optimization is posed as a summation problem with multiple observed data whereas this is not the case in the current optimization problem. The quantity to be estimated is the convergent value of u_{12} for a given set of inputs $\boldsymbol{x}$. When the inputs are random, then the coupling variable u_{12} is also random and its probability distribution needs to be calculated. This can be viewed similar to a statistical parameter estimation problem, and can be approached from the perspective of likelihood-based estimation, where the goal is to estimate the coupling variable u_{12}.

Recall that the likelihood function is defined as being proportional to the "probability of observing the given data" conditioned on the parameter to be estimated (u_{12} in this case). However, the problem of estimating the PDF of the coupling variable u_{12} in multi-disciplinary analysis is purely an uncertainty propagation problem and there is no "data" to calculate the likelihood function of u_{12}. Hence, the definition of the likelihood function cannot be used directly.

However, the focus of the MDA problem is to satisfy the inter-disciplinary compatibility condition. Consider "the probability of satisfying the inter-disciplinary compatibility conditioned on u_{12}", which can be written as $P(U_{12} = u_{12}|u_{12})$. This definition is similar to the original definition of the likelihood function (Section 2.4.3). It is a weight that is associated with a particular value of u_{12} to satisfy the multi-disciplinary constraint. In other words, if the ratio of $P(U_{12} = u_{12}^{(1)}|u_{12}^{(1)})$ to $P(U_{12} = u_{12}^{(2)}|u_{12}^{(2)})$ is equal to 0.1, then it is 10 ten times more likely for $u_{12}^{(2)}$ than $u_{12}^{(1)}$ to satisfy the inter-disciplinary compatibility condition. Thus, the properties of this expression are similar to the properties of the original likelihood function (Section 2.4.3). Hence, this expression is defined to be the likelihood of u_{12} in this chapter, as shown in Eq. 7.2.

Since the likelihood function is meaningful only up to a proportionality constant, Eq. 7.2 also uses only a proportionality sign.

$$L(u_{12}) \propto P(U_{12} = u_{12}|u_{12}) \tag{7.2}$$

Note that this definition is in terms of probability and hence the tool of likelihood gives a systematic procedure for including the uncertainty in the inputs during the construction of likelihood and estimating the probability distribution of the coupling variables, as explained below.

Note that there is a convergent value of u_{12} for every realization of $\boldsymbol{x}$. If $\boldsymbol{x}$ is represented using a probability distribution, then one sample of $\boldsymbol{x}$ has a relative likelihood of occurrence with respect to another sample of $\boldsymbol{x}$. Correspondingly, a given sample of u_{12} has a relative likelihood of being a convergent solution with respect another sample of u_{12}, and hence u_{12} can be represented using a probability distribution. It is this likelihood function and the corresponding probability distribution that will be calculated using the proposed method.

For a given value of u_{12}, consider the operation $U_{12} = G(u_{12}, \boldsymbol{x})$ defined earlier in Eq. 7.1. When $\boldsymbol{x}$ is random, an uncertainty propagation method can be used to calculate the distribution of U_{12}. Let the probability density function of U_{12} be denoted by $f_{U_{12}}(U_{12}|u_{12})$.

The aim is to calculate the likelihood of u_{12}, i.e. $L(u_{12})$ as the probability of satisfying the multi-disciplinary constraint, i.e. $U_{12} = u_{12}$. Since $f_{U_{12}}(U_{12}|u_{12})$ is a continuous PDF, the probability that U_{12} is equal to any particular value, u_{12} in this case, is equal to zero. Pawitan [22] explained that this problem can be overcome by considering an infinitesimally small window $[u_{12} - \frac{\epsilon}{2}, u_{12} + \frac{\epsilon}{2}]$ around u_{12} by

acknowledging that there is only limited precision in the real world.

$$L(u_{12}) \propto P(U_{12} = u_{12}|u_{12}) = \int_{u_{12}-\frac{\epsilon}{2}}^{u_{12}+\frac{\epsilon}{2}} f_{U_{12}}(U_{12}|u_{12})dU_{12} \propto f_{U_{12}}(U_{12} = u_{12}|u_{12}) \quad (7.3)$$

Note that this equation is similar to the common practice of estimating the parameters of a probability distribution given observed data for the random variable (Eq. 3.1 in Section 3.4).

Note that the likelihood function $L(u_{12})$ is conditioned on u_{12} and hence the PDF of U_{12} is always conditioned on u_{12}. Once the likelihood function of u_{12}, i.e the probability of satisfying the multi-disciplinary compatibility for a given value of u_{12}. is calculated, the PDF of the converged value of the coupling variable u_{12} can be calculated as:

$$f(u_{12}) = \frac{L(u_{12})}{\int L(u_{12})du_{12}} \quad (7.4)$$

In the above equation, the domain of integration for the variable u_{12} is such that $L(u_{12}) \neq 0$. Note that Eq. 7.4 is a form of Bayes theorem with a non-informative uniform prior density for u_{12}. Once the PDF of u_{12} is calculated, the multi-disciplinary analysis with uni-directional coupling in Fig. 7.3 can be used in lieu of the multi-disciplinary analysis with bi-directional coupling in Fig. 7.1. The system output f can then be calculated using well-known methods of uncertainty propagation such as Monte Carlo sampling (MCS), first-order reliability method (FORM), second-order reliability method (SORM), etc.

During the aforementioned uncertainty propagation, the converged u_{12} and $\boldsymbol{x}$ are considered as independent inputs in order to compute the uncertainty in u_{21}, g_1, g_2, and f. However, for every given value of $\boldsymbol{x}$, there is only one converged value of u_{12}; this is not a statistical dependence but a functional dependence. The functional dependence between the converged u_{12} and $\boldsymbol{x}$ is not known and not considered in

the decoupled approach. If this functional dependence needs to be explicitly considered, one would have to revert to the computationally expensive fixed point iteration approach for every sample of $\boldsymbol{x}$. (An alternative would be to choose a few samples of $\boldsymbol{x}$, run fixed point iteration analysis on each of them and construct a surrogate/approximation of the functional dependence between $\boldsymbol{x}$ and u_{12}, and explicitly use this surrogate in uncertainty propagation. Obviously, the surrogate could also be directly constructed for any of the responses - g_1, g_2, or f - instead of considering the coupling variable u_{12}. However, replacing the entire multi-disciplinary analysis by a surrogate model is a different approach and does not fall within the scope of the decoupled approach).

The above discussion calculated the PDF of u_{12} and cut the coupling from "Analysis 1" to "Analysis 2". Without loss of generality, the same approach can be used to calculate the PDF of u_{21} and cut the coupling from "Analysis 2" to "Analysis 1". This method has several advantages:

1. This method is free from first-order or second-order approximations of the coupling variables.

2. The equations of the individual disciplinary analyses are not approximated during the derivation of Eq. 7.3 and the calculation of the PDF of the coupling variables in Eq. 7.4 is exact from a theoretical perspective.

3. The method does not require any coupled system analysis, i.e. repeated iteration between "Analysis 1" and "Analysis 2" until convergence.

Though the computation of the PDF of u_{12} is theoretically exact, two issues need to be addressed in computational implementation. (1) The calculation of $L(u_{12})$ requires the estimation of $f_{U_{12}}(U_{12}|u_{12})$ which needs to be calculated by propagating the inputs $\boldsymbol{x}$ through G for a given value of u_{12}. (2) This likelihood function needs

to be calculated for several values of u_{12} to perform the integration in Eq. 7.4. These two steps, i.e. uncertainty propagation and integration, could make the methodology computationally expensive if a Monte Carlo-type approach is pursued for uncertainty propagation.

Therefore, the following section proposes a methodology that makes the numerical implementation inexpensive for the above two steps. From here on, there are approximations made; note that these approximations are only for the purpose of numerical implementation and not a part of the mathematical theory. Here, "theory" refers to the development and use of Eq. 7.3 and Eq. 7.4 for uncertainty quantification in multi-disciplinary analysis, and "implementation" refers to the numerical computation of $f_{U_{12}}(U_{12} = u_{12}|u_{12})$ in Eq. 7.3.

7.4 Numerical Implementation

This section addresses the two issues mentioned above in the numerical implementation of the proposed likelihood-based approach.

7.4.1 Evaluation of the Likelihood Function $L(u_{12})$

The first task is to calculate the likelihood function $L(u_{12})$ for a given value of u_{12}. This requires the calculation of the PDF $f_{U_{12}}(U_{12}|u_{12})$. However it is not necessary to calculate the entire PDF. Based on Eq. 7.3, the calculation of likelihood $L(u_{12})$ only requires the evaluation of the PDF at u_{12}, i.e. $f_{U_{12}}(U_{12} = u_{12}|u_{12})$. Hence, instead of entirely evaluating the PDF $f_{U_{12}}(U_{12}|u_{12})$, only local analysis at $U_{12} = u_{12}$ needs to be performed. One method is to make use of FORM to evaluate this PDF value. This is the first approximation.

The first-order reliability method estimates the probability that a performance function $H = h(\boldsymbol{x})$ is less than or equal to zero, given uncertain input variables $\boldsymbol{x}$.

This probability is equal to the cumulative probability density (CDF) of the variable H evaluated at zero [27]. In this approach, the so-called most probable point (MPP) is calculated by transforming the variables $\boldsymbol{x}$ into uncorrelated standard normal space u and by determining the point in the transformed space that is closest to the origin. An optimization problem can be formulated as shown in Fig. 7.4.

Given PDFs of $\boldsymbol{x}$
Minimize $\beta = \boldsymbol{u}^T \boldsymbol{u}$
such that $H \equiv h(\boldsymbol{x}) = 0$
where standard normal $\boldsymbol{u} = T(\boldsymbol{x})$
$P(H \leq 0) = \Phi(-\beta)$

Figure 7.4: Use of FORM to Estimate the CDF Value

The details of the transformation $u = T(x)$ in Fig. 7.4 can be found in Haldar and Mahadevan [27]. This optimization can be solved by using the well-known Rackwitz-Fiessler algorithm [31], which is based on a repeated linear approximation of the constraint $H = 0$. Once the shortest distance to the origin is estimated to be equal to β, then the CDF value is calculated in FORM as:

$$P(H \leq 0) = \phi(-\beta) \tag{7.5}$$

FORM can also be used to calculate the CDF value at any generic value h_c, i.e. $P(h(\boldsymbol{x}) \leq h_c)$ and the probability that $h(\boldsymbol{x})$ is less than or equal to h_c can be evaluated by executing the FORM analysis for the performance function $H = h(\boldsymbol{x}) - h_c$. For the problem at hand, it is necessary to calculate the PDF value at u_{12} and not the CDF value. This can be accomplished by finite differencing, i.e. by performing two FORM analyses at $h_c = u_{12}$ and $h_c = u_{12} + \delta$, where δ is a small difference that can be chosen, for example, $0.001 \times u_{12}$. The resultant CDF values from the two FORM

analyses are differenced and divided by δ to provide an approximate value of the PDF value at u_{12}. This is the second approximation.

Hence, the evaluation of the likelihood function $L(u_{12})$ is based on two approximations: (1) the PDF value is calculated based on finite differencing two CDF values; and (2) each CDF value is in turn calculated using FORM which is a first-order approximation (Eq. 7.5).

7.4.2 Construction of PDF of u_{12}

Recall that Eq. 7.4 is used to calculate the PDF of u_{12} based on the likelihood function $L(u_{12})$. In theory, for any chosen value of u_{12}, the corresponding likelihood $L(u_{12})$ can be evaluated, and hence the integral in Eq. 7.4 can be computed. For the purpose of numerical implementation, the limits of integration need to be chosen. The first-order estimates of the mean and variance of u_{12} can be estimated by calculating the converged value of u_{12} at the mean of the uncertain input values using fixed point iteration. The derivatives of the coupling variables with respect to the inputs can be calculated using Sobieski's system sensitivity equations [229], as demonstrated later in Section 7.4.1. These first order estimates can be then used to select the limits (for example, six sigma limits) for integration.

In order to evaluate the PDF of u_{12}, The likelihood function is evaluated only at a few points; a recursive adaptive version of Simpson's quadrature [132] is used to evaluate this integral and the points at which the likelihood function needs to be evaluated are adaptively chosen until the quadrature algorithm converges.

This quadrature algorithm is usually applicable only in the case of one-dimensional integrals whereas in a typical multi-disciplinary problem, u_{12} may be a vector, where there are several coupling variables in each direction. Hence, the multi-dimensional integral is decomposed into multiple one-dimensional integrals so that the quadrature

algorithm may be applied.

$$\int L(\alpha,\beta)d\alpha d\beta = \int \Big(\int L(\alpha,\beta)d\alpha \Big) d\beta \tag{7.6}$$

Each one-dimensional integral is evaluated using recursive adaptive Simpson's quadrature algorithm [132], which was explained earlier in Section 4.3.7.

This technique ensures that the number of evaluations of the individual disciplinary analyses is minimal. Would it be possible to approximately estimate the number of disciplinary analyses needed for uncertainty propagation? Suppose that the likelihood function is evaluated at ten points to solve the integration in Eq. 7.4. Each likelihood evaluation requires a PDF calculation, and hence two FORM analyses. Assume that the optimization for FORM converges in five iterations on average; each iteration would require $\underline{n+1}$ (where n is the number of input variables) evaluations of the individual disciplinary analysis (one evaluation for the function value and n evaluations for derivatives). Thus, the number of individual disciplinary analyses required will approximately be equal to $100(n+1)$. This is computationally efficient when compared to existing approaches. For example, Mahadevan and Smith [222] report that for a multi-disciplinary analysis with 5 input variables, the multi-constraint FORM approach required 69 evaluations for the evaluation of a single CDF value, which on average may lead to 690 evaluations for 10 CDF values. While the proposed method directly calculates the entire PDF, it also retains the functional dependence between the disciplinary analyses, thereby enabling uncertainty propagation to the next analysis level.

As the number of coupling variables increases, the integration procedure causes the computational cost to increase exponentially. For example, if there are ten coupling variables, each with 5 discretization points (for the sake of integration), then the number of individual disciplinary analyses required will approximately be equal to

$5^{10} \times 10 \times (n+1)$. Alternatively, a sampling technique such as Markov Chain Monte Carlo (MCMC) sampling can be used to draw samples of the coupling variables; this method can draw samples of the coupling variable without evaluating the integration constant in Eq. 7.4. Further, since this is sampling approach, the computational cost does not increase exponentially with the number of coupling variables. In each iteration of the MCMC chain, two FORM analyses need to be conducted to evaluate the likelihood for a given value of u_{12} (which is now vector), and several thousands (say, Q) of evaluations of this likelihood function may be necessary for generating the entire PDFs of the coupling variables. Thus, the number of individual disciplinary analyses will be approximately equal to $10 \times (n+1) \times Q$. Currently, the proposed method is demonstrated only for a small number of coupling variables. Future work will extend the methodology to field-type quantities (temperatures, pressures, etc. in finite element analysis) where the number of coupling variables is large.

7.5 Example: Mathematical MDA Problem

7.5.1 Description of the Problem

This problem consists of three analyses, two of which are coupled with one another. This is an extension of the problem discussed by Du and Chen [221], and later by Mahadevan and Smith [222] where only two analyses where considered. The functional relationships are shown in Fig. 7.5.

In addition to the two analyses given in [222], the current study considers a third analysis where a system output is calculated based on g_1 and g_2 as $f = g_2 - g_1$. All the five input quantities $\boldsymbol{x} = (x_1, x_2, x_3, x_4, x_5)$ are assumed to be normally distributed (only for the sake of illustration) with unit mean and standard deviation equal to 0.1; there is no correlation between them. The goal in [221] and [222] was to calculate the

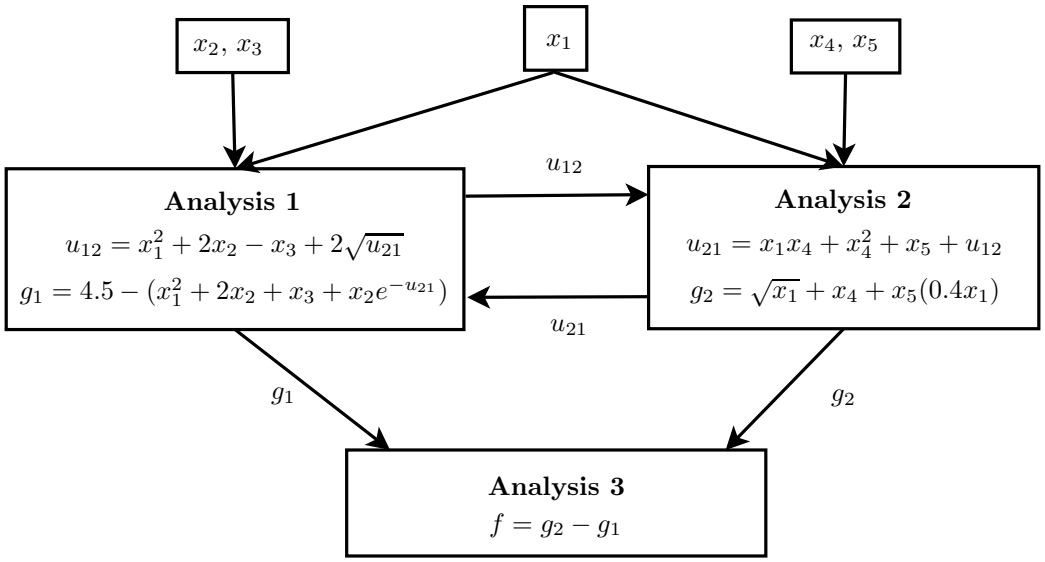


Figure 7.5: Functional Relationships

probability $P(g_1 \leq 0)$. In this section, the goal is to calculate the entire probability distributions of the coupling variables u_{12} and u_{21}, the outputs of the individual analyses g_1 and g_2, and the overall system output f.

A coarse approximation of the uncertainty in the output variables and coupling variables can be obtained in terms of first-order mean and variance using Taylors series expansion [27]. For example, consider the coupling variable u_{12}; the procedure described for can be extended to u_{21}, g_1, g_2, and f. The first-order mean of u_{12} can be estimated by calculating the converged value of u_{12} at the mean of the input values, i.e. $\boldsymbol{x} = (1, 1, 1, 1, 1)$. The first-order mean values of of u_{12}, u_{21}, g_1, g_2, and f are calculated to be equal to 8.9, 11.9, 0.5, 2.4, and 1.9 respectively. The first-order variance of u_{12} can be estimated as:

$$Var(u_{12}) = \sum_{i=1}^{n} \left(\frac{du_{12}}{dx_i}\right)^2 Var(x_i) \tag{7.7}$$

where the first-order derivatives are calculated using Sobieski's system (or global)

sensitivity equations [229], by satisfying the multi-disciplinary compatibility as:

$$\frac{du_{12}}{dx_i} = \frac{\partial u_{12}}{\partial x_i} + \frac{\partial u_{12}}{\partial u_{21}} \frac{\partial u_{21}}{\partial x_i} \tag{7.8}$$

All the derivatives are calculated at the mean of the input values, i.e. $\boldsymbol{x} = (1, 1, 1, 1, 1)$. The values of $\frac{\partial u_{12}}{\partial x_i}$ are 2, 2, -1, 0 and 0 ($i = 1$ to 5) respectively. The values of $\frac{\partial u_{21}}{\partial x_i}$ are 1, 0, 0, 3, and 1 ($i = 1$ to 5) respectively. The value of $\frac{\partial u_{12}}{\partial u_{21}}$ is $\frac{1}{\sqrt{u_{21}}}$, evaluated at the mean, which is equal to 0.29. Hence, using Eq. 7.7 and Eq. 7.8, the standard deviation of u_{12} is calculated to be 0.333.

The system sensitivity equation-based approach only provides approximations of the mean and variance, and it cannot calculate the entire PDF of u_{12}. The remainder of this section illustrates the proposed LAMDA approach, which can accurately calculate the entire PDF of u_{12}. Though the system of equations in Fig. 7.5 may be solved algebraically by eliminating one variable, the numerical implementation does not take advantage of this closed form solution and assumes each analysis to be a black-box. This is done to simulate the behavior of realistic multi-disciplinary analyses that may not have closed form solutions. For the same reason, finite differencing is used to approximate the gradients even though analytical derivatives can be calculated easily for this problem.

7.5.2 Calculation of the PDF of the Coupling Variable

In this numerical example, the coupling variable u_{12} is estimated for the sake of illustration, and the arrow from "Analysis 1" to "Analysis 2" is severed. The PDF of u_{12} is estimated using the proposed methods (1) sampling with optimization-based deterministic MDA (SOMDA); and (2) likelihood approach for multi-disciplinary analysis (LAMDA). In Fig. 7.6, the PDF using the LAMDA method uses 10 integration points for the evaluation of Eq. 7.4. The resulting PDFs from the SOMDA method and the

LAMDA method are compared with the benchmark solution which is estimated using $10,000$ Monte Carlo samples of $\boldsymbol{x}$ and fixed point iteration (until convergence of Analysis 1 and Analysis 2) for each sample of $\boldsymbol{x}$. The probability bounds on MCS results for the benchmark solution are also calculated using the formula $CoV(F) = \sqrt{\frac{(1-F)}{F.N}}$ where F is the CDF value [27], and found to be narrow and almost indistinguishable from the solution reported in Fig. 7.6. Since the benchmark solution uses fixed point iteration (FPI) for each input sample, it is indicated as SOFPI (sampling outside fixed point iteration) in Fig. 7.6.

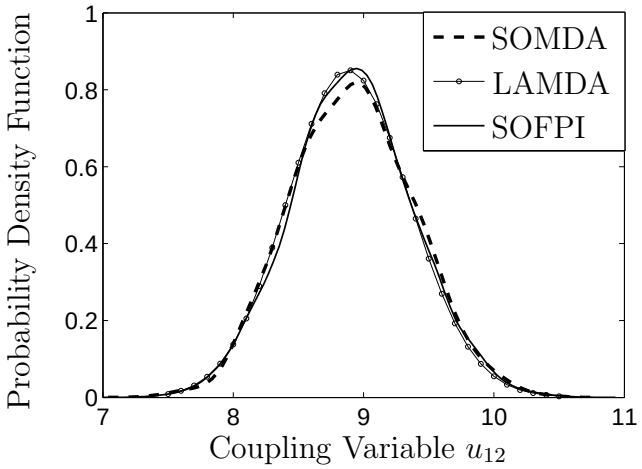
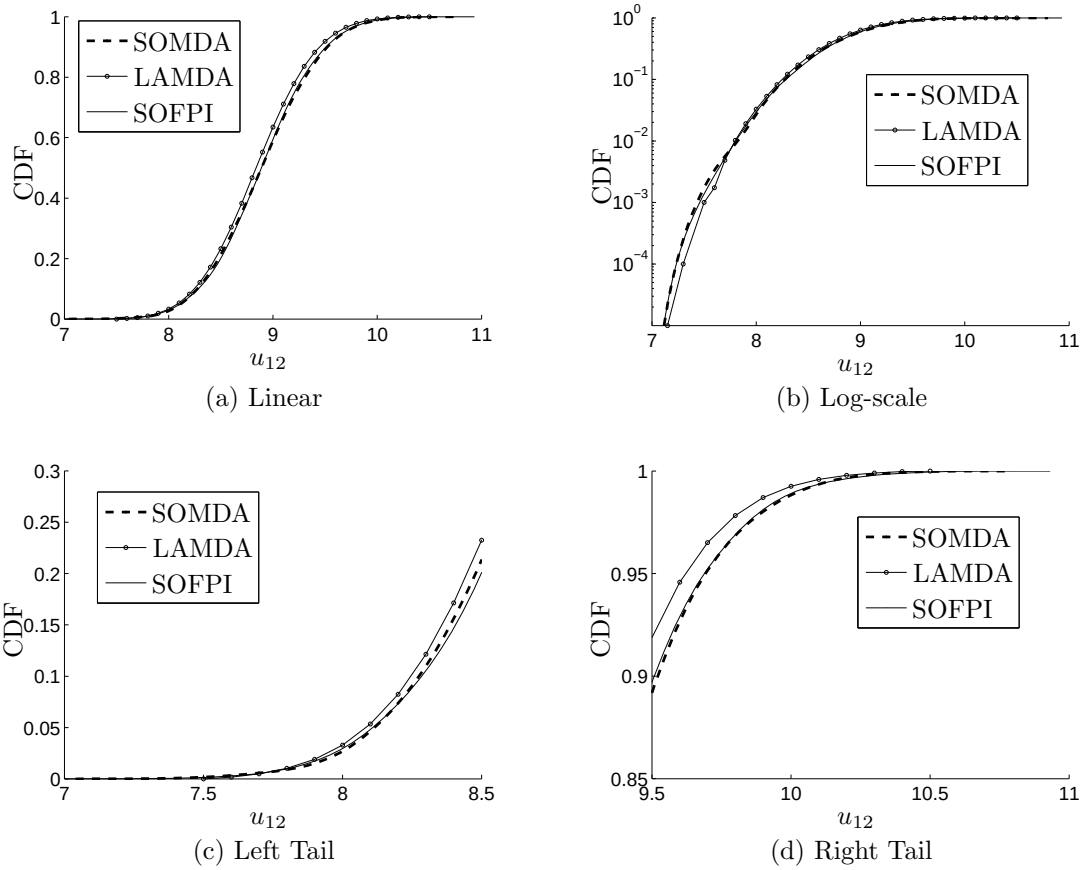


Figure 7.6: PDF of u_{12}

In addition to the PDF, the CDF of u_{12} is shown in Fig. 7.6. The CDF is plotted in linear and log-scale. Further, the tail probabilities are important in the context of reliability analysis; hence, the two tails of the CDF curves are also shown separately.

It is seen that the solutions (PDF values and CDF values) from the proposed method (LAMDA) match well with the benchmark (SOFPI) solution and the SOMDA approach. Note that the mean and standard deviation of the PDF in Fig. 7.6 agree well with the first-order approximations previously calculated (8.9 and 0.333). Obviously, the above PDF provides more information than the first-order mean and



(a) Linear (b) Log-scale

(c) Left Tail (d) Right Tail

Figure 7.7: Cumulative Distribution Function of u_{12}

standard deviation, and is more suitable for calculation of tail probabilities in reliability analysis.

The differences (maximum error is less than 1%) seen in the PDFs and the CDFs from the three methods, though small, are accountable. The PDF obtained using SOMDA differs from the benchmark solution because it uses only 1000 Latin hypercube samples (realizations of inputs) whereas the benchmark solution used $10,000$ samples. The PDF obtained using LAMDA differs from the benchmark solution because of two approximations (1) finite differencing two CDF values to calculate the PDF value, and (2) calculating each CDF value using FORM.

The benchmark solution is based on fixed point iteration and required about 10^5

evaluations each of Analysis 1 and Analysis 2. The SOMDA method required 8000 - 9000 executions of each individual disciplinary analysis. (This number depends on the random samples of the input, since for each sample, the number of optimization iterations required for convergence is different). Note that theoretically, the SOMDA method would produce a PDF that is identical to the benchmark solution if the same set of input samples were used in both the cases. This is because the SOMDA approach simply solves the deterministic MDA problem and then considers sampling in an outside loop. The solution approach in SOMDA is different from that in the benchmark solution approach; however, the treatment of uncertainty is the same. As discussed in Section 7.2, the SOMDA method is still expensive; replacing the brute force fixed point iteration in the benchmark solution by an optimization did not significantly improve the computational efficiency in this problem.

The LAMDA method treats the uncertainty directly in the definition of likelihood, and was found to be the least expensive, as it required only about 450 - 500 evaluations of each disciplinary analysis for the estimation of the entire PDF of u_{12} in Fig. 7.6. The number of evaluations is given as a range because of three sources of variation: (1) different initial guesses for FORM analyses may require different numbers of function evaluations for convergence to MPP; (2) the number of integration points used for evaluation of Eq. 7.4; and (3) the actual values of the integration points used for evaluation of Eq. 7.4. In contrast, the multi-constraint FORM approach developed by Mahadevan and Smith [222] required about 69 evaluations for the calculation of the CDF at one particular value. If the entire PDF as in Fig. 7.6 is desired, the multi-constraint FORM would take approximately $69 \times 2n$ function evaluations, where n is the number of points on the PDF and each PDF evaluation would require 2 CDF evaluations.

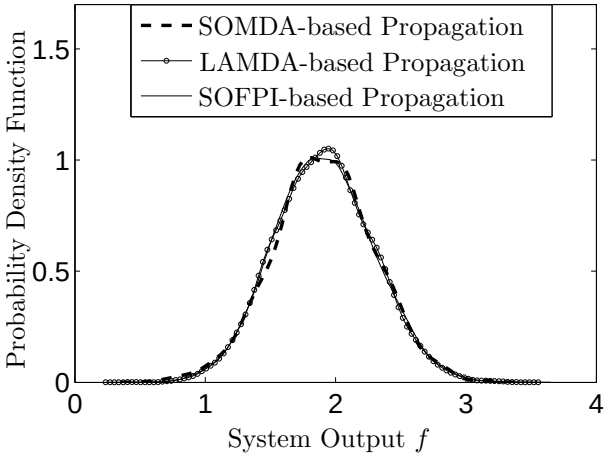


Figure 7.8: PDF of f

7.5.3 Calculation of PDF of the System Output

Once the PDF of u_{12} is calculated, the scheme in Fig. 7.3 can be used for uncertainty propagation and the PDF of the system output f is calculated. Note that this does not require any multi-disciplinary analysis (iterative analysis between the two subsystems) and it is now a simple uncertainty propagation problem. Well-known methods for uncertainty propagation such as Monte Carlo Simulation (MCS), First Order and Second Order Reliability Methods (FORM, SORM) [27] can be used for this purpose. For the sake of illustration, Monte Carlo simulation is used. The PDF of the system output f is shown in Fig. 7.8.

As the coupling variable u_{12} has been estimated here, the "arrow" from Analysis 1 to Analysis 2 alone is severed, whereas the arrow from Analysis 2 to Analysis 1 is retained. Hence, to solve for the system output f, the probability distributions of the inputs $\boldsymbol{x}$ and the probability distribution of the coupling variable u_{12} are used first in Analysis 2 (to calculate u_{21}), and then in Analysis 1 to calculate the individual disciplinary system outputs g_1 and g_2, followed by the overall system output f. As seen from Fig. 7.8 , the solutions from the three different methods sampling

with optimization-based deterministic MDA (SOMDA), likelihood approach for multi-disciplinary analysis (LAMDA), and the benchmark solution (SOFPI) compare well against each other. The following section uses the proposed methodology in the analysis of an aerospace application, a satellite used to detect forest fires.

7.6 Three-discipline Fire Detection Satellite Model

This section first describes the components of the satellite system model used for detection of forest fires, and then illustrates how this problem is representative of a multi-disciplinary problem. Then, the proposed methods are used for multi-disciplinary uncertainty propagation analysis.

7.6.1 Description of the Problem

This problem was originally described by Wertz and Larson [230]. This is a hypothetical but realistic spacecraft consisting of a large number of subsystems with both feedback and feed-forward couplings. The primary objective of this satellite is to detect, identify, and monitor forest fires in near real time. This satellite is intended to carry a large and accurate optical sensor of length 3.2 m, weight 720 kg and has an angular resolution of 8.8×10^{-7} radians. This section uses the modified version of this problem considered earlier by Ferson et al. [231] and Zaman [228].

Zaman [228] considered a subset of three subsystems of the fire detection satellite, consisting of i) Orbit Analysis, ii) Attitude Control and iii) Power, based on Ferson et al [231]. This three-subsystem problem is shown in Fig. 7.9. There are nine random variables in this problem, as indicated in Fig. 7.9.

As seen in Fig. 7.9, the Orbit subsystem has feed-forward coupling with both Attitude Control and Power subsystems, whereas the Attitude Control and Power subsystems have feedback or bi-directional coupling through three variables P_{ACS},

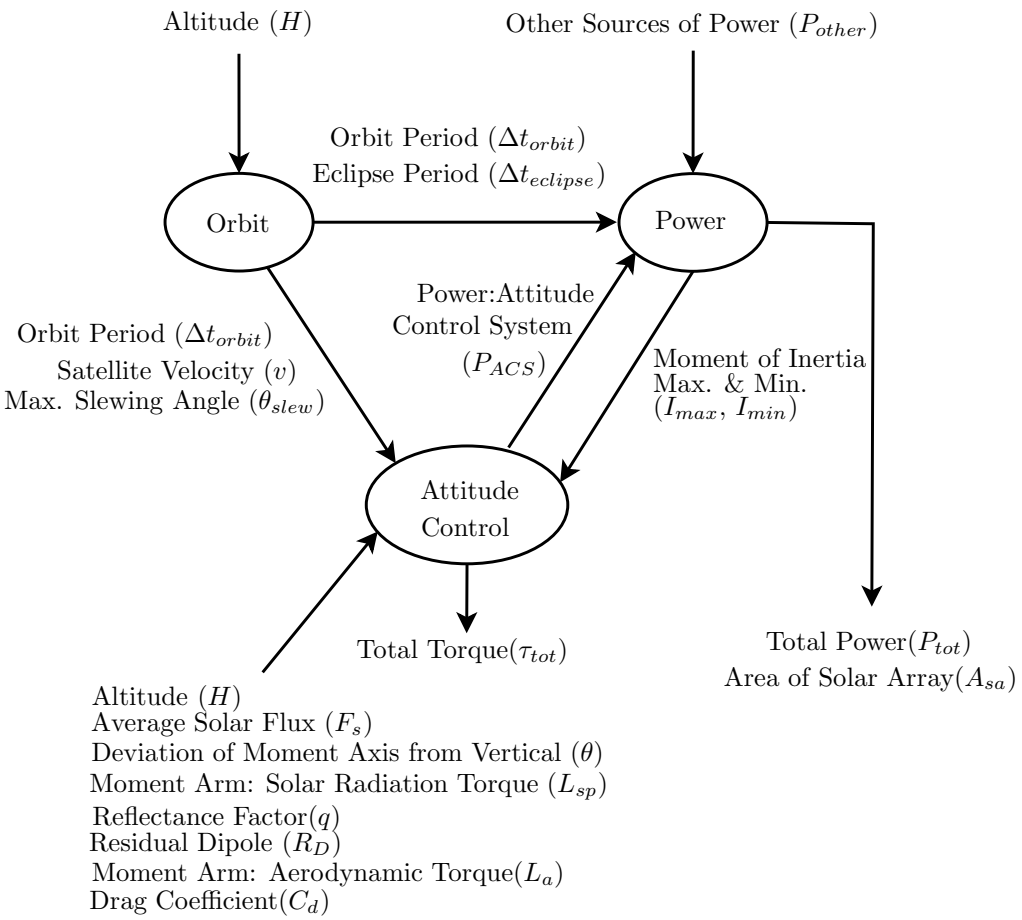


Figure 7.9: A Three Subsystem Fire Detection Satellite

I_{min}, and I_{max}. A satellite configuration is assumed in which two solar panels extend out from the spacecraft body. Each solar panel has dimensions L by W and the inner edge of the solar panel is at a distance D from the centerline of the satellites body as shown in Fig. 7.10.

The functional relationships between the three subsystems are developed in detail by Wertz and Larson [230] and summarized by Ferson et al. [231]. These functional relationships are briefly described in this section.

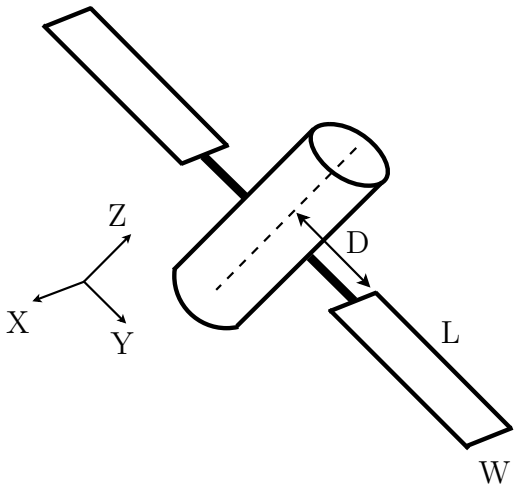


Figure 7.10: Schematic Diagram for the Spacecraft Solar Array

7.6.1.1 The Orbit Subsystem

The inputs to this subsystem are: radius of the earth (R_E); orbit altitude (H); earths standard gravitational parameter (μ); and target diameter (ϕ_{target}).

The outputs of this subsystem are: satellite velocity (v); orbit period (Δt_{orbit}); eclipse period ($\Delta t_{eclipse}$); and maximum slewing angle (θ_{slew}). The relationships between these variables are summarized in the following equations:

$$v = \sqrt{\frac{\mu}{R_E + H}} \tag{7.9}$$

$$\Delta t_{orbit} = 2\pi\sqrt{\frac{(R_E + H)^3}{\mu}} = \frac{2\pi(R_E + H)}{v} \tag{7.10}$$

$$\Delta t_{eclipse} = \frac{\Delta t_{orbit}}{\pi} \arcsin\left(\frac{R_E}{R_E + H}\right) \tag{7.11}$$

$$\theta_{slew} = \arctan\left(\frac{\sin\left(\frac{\phi_{target}}{R_E}\right)}{1 - \cos(\frac{\phi_{target}}{R_E}) + \frac{H}{R_E}}\right) \tag{7.12}$$

7.6.1.2 The Attitude Control Subsystem

The 23 inputs to this subsystem are: earths standard gravitational parameter (μ); radius of the earth (R_E); Altitude (H); maximum and minimum moment of inertia of the spacecraft (I_{max} and I_{min}); deviation of major moment axis from local vertical (θ); moment arm for the solar radiation torque (L_{sp}); average solar flux (F_s); speed of light (c); reflectance factor (q); surface area off which solar radiation is reflected (A_s); Slewing time period (Δt_{slew}); magnetic moment of the Earth (M); residual dipole of the spacecraft (R_D); moment arm for aerodynamic torque (L_a); atmospheric density (ρ); maximum slewing angle (θ_{slew}); sun incidence angle (i); drag coefficient (C_d); cross sectional surface area in the direction of flight (A); satellite velocity (v); rotation velocity of reaction wheel (ω_{max}); number of reaction wheels (n); and holding power (P_{hold}), i.e. the power required to maintain the constant velocity (ω_{max}). The overall output of this subsystem is the total torque (τ_{tot}). The value of the total torque is computed based on slewing torque (τ_{slew}), disturbance torque (τ_{dist}), gravity gradient torque (τ_g), solar radiation torque (τ_{sp}), magnetic field interaction torque (τ_m), and aerodynamic torque (τ_a), as shown in the following equations.

$$\tau_{tot} = max(\tau_{slew}, \tau_{dist}) \tag{7.13}$$

$$\tau_{slew} = \frac{4\theta_{slew}}{(\Delta t_{slew})^2} I_{max} \tag{7.14}$$

$$\tau_{dist} = \sqrt{\tau_g^2 + \tau_{sp}^2 + \tau_m^2 + \tau_a^2} \tag{7.15}$$

$$\tau_g = \frac{3\mu}{2(R_E + H)^3} |I_{max} - I_{min}| \sin(2\theta) \tag{7.16}$$

$$\tau_{sp} = L_{sp} \frac{F_s}{C} A_s (1 + q) \cos(i) \tag{7.17}$$

$$\tau_m = \frac{2MR_D}{(R_E + H)^3} \tag{7.18}$$

$$\tau_a = \frac{1}{2} L_a \rho C_d A v^2 \tag{7.19}$$

Note that this subsystem takes two coupling variables (I_{max} and I_{min}) as input and produces another coupling variable (Attitude control power: P_{ACS}) as output, as given in the following equation.

$$P_{ACS} = \tau_{tot} \omega_{max} + n P_{hold} \tag{7.20}$$

This coupling variable is an input to the power subsystem, as described in the following subsection.

7.6.1.3 The Power Subsystem

The 16 inputs to the power subsystem are: attitude control power (P_{ACS}); other sources of power (P_{other}); orbit period (Δt_{orbit}); eclipse period ($\Delta t_{eclipse}$); sun incidence angle (i); inherent degradation of the array (I_d); average solar flux (F_s); power efficiency (η); lifetime of the spacecraft (L_T); degradation in power production capacity in % per year (ϵ_{deg}); length to width ratio of solar array (r_{lw}); number of solar arrays (n_{sa}); average mass density of solar arrays (ρ_{sa}); thickness of solar panels (t); distance between the panels (D); and moments of inertia of the main body of the spacecraft (I_{bodyX}, I_{bodyY}, I_{bodyZ}).

The overall outputs of this subsystem are the total power (P_{tot}), and the total size of the solar array (A_{sa}), as calculated below.

$$P_{tot} = P_{ACS} + P_{other} \tag{7.21}$$

Let P_e and P_d denote the spacecrafts power requirements during eclipse and daylight, respectively. For the sake of illustration, it is assumed that $P_e = P_d = P_{tot}$. Let T_e and T_d denote the time per orbit spent in eclipse and in sunlight, respectively. It is

assumed that $Te = \Delta t_{eclipse}$ and $T_d = \Delta t_{orbit} - T_e$. Then the required power output (P_{sa}) is calculated as:

$$P_{sa} = \frac{(\frac{P_e T_e}{0.6} + \frac{P_d T_d}{0.8})}{T_d} \tag{7.22}$$

The power production capabilities at the beginning of life (P_{BOL}) and at the end of the life (P_{EOL}) are calculated as:

$$\begin{aligned} P_{BOL} &= \eta F_s I_d \cos(i) \\ P_{EOL} &= P_{BOL}(1 - \epsilon_{deg})^{LT} \end{aligned} \tag{7.23}$$

The total solar array size, i.e. the second output of this subsystem, is calculated as:

$$A_{sa} = \frac{P_{sa}}{P_{EOL}} \tag{7.24}$$

Note that this subsystem takes a coupling variable (P_{ACS}) as input and produces the other two coupling variables (I_{max} and I_{min}) as output, to be fed into the attitude control subsystem described earlier.

The length (L), width (W), mass (m_{sa}), moments of inertia (I_{saX}, I_{saY}, I_{saZ}) of the solar array are calculated as follows:

$$\begin{aligned} L &= \sqrt{\frac{A_{sa} r_{lw}}{m_{sa}}} \\ W &= \sqrt{\frac{A_{sa}}{r_{lw} m_{sa}}} \\ m_{sa} &= 2\rho_{sa} LWt \end{aligned} \tag{7.25}$$

$$I_{saX} = m_{sa}\left[\frac{1}{12}(L^2 + t^2) + \left(D + \frac{L}{2}\right)^2\right] \tag{7.26}$$

$$I_{saY} = \frac{m_{sa}}{12}(t^2 + W^2) \tag{7.27}$$

$$I_{saZ} = m_{sa}\left[\frac{1}{12}(L^2 + W^2) + \left(D + \frac{L}{2}\right)^2\right] \tag{7.28}$$

The total moment of inertia (I_{tot}) can be computed in all three directions (X, Y, and Z), from which the maximum and the minimum moments of inertia (I_{max} and I_{min}) can be computed.

$$I_{tot} = I_{sa} + I_{body} \tag{7.29}$$

$$I_{max} = max(I_{totX}, I_{totY}, I_{totZ}) \tag{7.30}$$

$$I_{min} = min(I_{totX}, I_{totY}, I_{totZ}) \tag{7.31}$$

7.6.2 Numerical Details

Some of the input quantities are chosen to be stochastic while others are chosen to be deterministic. Table 7.1 provides the numerical details for the deterministic quantities and Table 7.2 provides the numerical details for the stochastic quantities. All the stochastic quantities are treated to be normally distributed, for the sake of illustration.

7.6.3 Uncertainty Propagation Problem

As seen in Fig. 7.9, this is a 3-disciplinary analysis problem, with feedback coupling between two disciplines "power" and "attitude control". It is required to compute the uncertainty in 3 system output variables total power P_{tot}, required solar array area A_{sa}, and total torque τ_{tot}.

Prior to the quantification of the outputs, the first step is the calculation of the probability distribution of the coupling variables. The functional dependency can be severed in either direction, either from "power"to "attitude control" or from "attitude control" to "power", and this choice can be made without loss of generality. In this example, the probability distribution of P_{ACS}, i.e. the power of the attitude control system is calculated; this distribution now becomes an independent input

Table 7.1: List of Deterministic Quantities

Variable	Symbol	Unit	Numerical Value
Earth's Radius	R_E	m	6378140
Gravitational Parameter	μ	$m^3 s^{-2}$	3.986×10^{14}
Target Diameter	ϕ_{target}	m	235000
Light Speed	c	ms^{-1}	2.9979×10^8
Area Reflecting Radiation	A_s	m^2	13.85
Sun Incidence Angle	i	degree	0
Slewing Time Period	Δt_{slew}	sec	760
Magnetic Moment of Earth	M	$A.m^2$	7.96×10^{15}
Atmospheric Density	ρ	kgm^{-3}	5.1480×10^{-11}
Cross-sectional in Flight Direction	A	m^2	13.85
No. of Reaction Wheels	n	–	3
Maximum Velocity of a Wheel	ω_{max}	rpm	6000
Holding Power	P_{hold}	W	20
Inherent Degradation of Array	I_d	-	0.77
Power Efficiency	η	-	0.22
Lifetime of Spacecraft	LT	years	15
Degradation in Power Production Capacity	ϵ_{deg}	% per year	0.0375
Length to Width Ratio of Solar Array	r_{lw}	-	3
Number of Solar Arrays	n_{sa}	-	3
Average Mass Density to Arrays	ρ_{sa}	kgm^3	700
Thickness of Solar Panels	t	m	0.005
Distance between Panels	D	m	2
Moments of Inertia of spacecraft body	I_{body}	kgm^2	$I_{body,X} = 6200$ $I_{body,Y} = 6200$ $I_{body,Z} = 4700$

to the "power subsystem"; the functional dependency between "power" to "attitude control" is retained through the two coupling variables in the opposite direction. The following subsections present these results; Section 7.6.4 calculates the PDF of the feedback variable P_{ACS} and Section 7.6.5 calculates the PDFs of the system outputs.

7.6.4 Calculation of PDF of the Coupling Variable

Similar to the mathematical example presented in Section 7.5, this section calculates the PDF of the coupling variable P_{ACS} using sampling with optimization-based deterministic MDA (SOMDA) and the likelihood approach for multi-disciplinary analysis (LAMDA). These results are compared with the benchmark solution in Fig. 7.11.

Table 7.2: List of Stochastic Quantities

Variable	Symbol	Unit	Mean	Standard Deviation
Altitude	H	m	18000000	1000000
Power other than ACS	P_{other}	W	1000	50
Average Solar Flux	F_s	W/m^2	1400	20
Deviation of Moment Axis	θ	degree	15	1
Moment Arm for Radiation Torque	L_{sp}	m	2	0.4
Reflectance Factor	q	-	0.5	1
Residual Dipole of Spacecraft	R_D	Am^2	5	1
Moment Arm for Aerodynamic Torque	L_a	m	2	0.4
Drag Coefficient	C_d	-	1	0.3

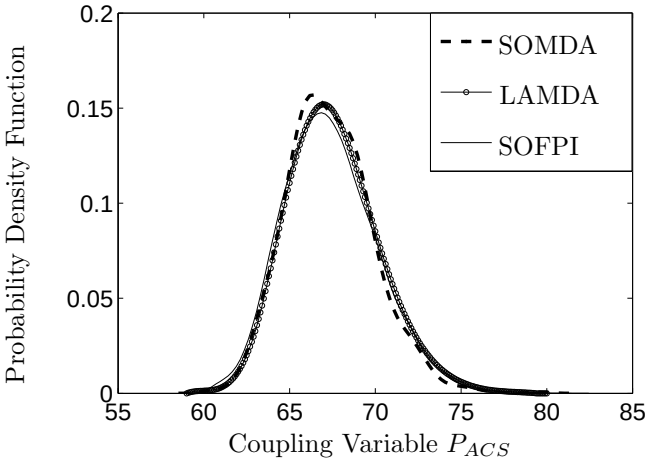


Figure 7.11: PDF of Coupling Variable P_{ACS}

In Fig. 7.11, the PDF using the LAMDA method uses 10 integration points for the evaluation of Eq. 7.4.

Similar to the mathematical example in Section 7.5, it is seen from Fig. 7.11 that the results from SOMDA and LAMDA compare well with the benchmark solution (SOFPI). In addition to the PDFs, the CDFs and the tail probabilities are also in reasonable agreement. The benchmark solution is based on fixed point iteration and required about $200,000$ evaluations each of the power subsystem and the attitude control subsystem. The SOMDA method required about 20000 evaluations whereas the LAMDA method required about 900- 1000 evaluations. It is clear that the LAMDA approach provides an efficient and accurate alternative to sampling-based approaches.

7.6.5 Calculation of PDFs of the System Outputs

Once the probability distribution of the coupling variable P_{ACS} is calculated, the system does not contain any feedback coupling and hence, methods for simple forward uncertainty propagation can be used to estimate the PDFs of the three system outputs total power (P_{tot}), required solar array area (A_{sa}), and total torque (τ_{tot}). Monte Carlo simulation is used for uncertainty propagation. The resulting PDFs are plotted in Figs. 7.12 - 7.14.

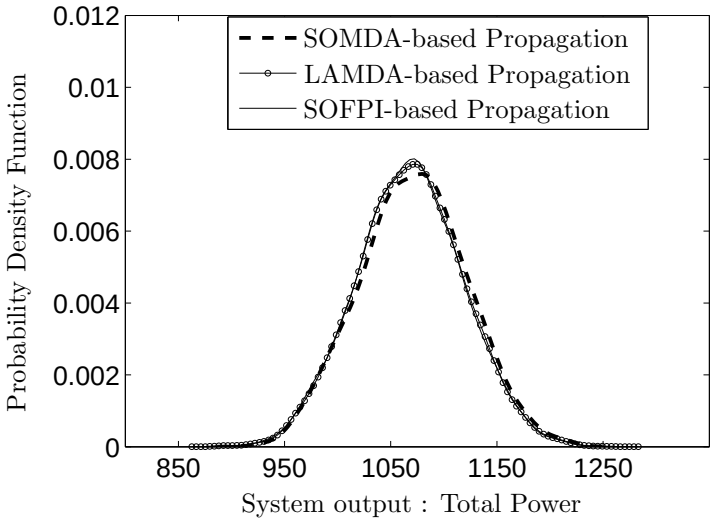


Figure 7.12: PDF of Total Output Power P_{tot}

As seen from Figs. 7.12 - 7.14, the PDFs of the system outputs obtained using both SOMDA and LAMDA compare well with the benchmark solution (SOFPI).

7.7 Conclusion

Existing methods for uncertainty propagation in multi-disciplinary system models are based on (1) Monte Carlo sampling around fixed point iteration, which is computationally expensive; and/or (2), approximating the system equations; and/or (3) approximating the probability distributions of the coupling variables and then

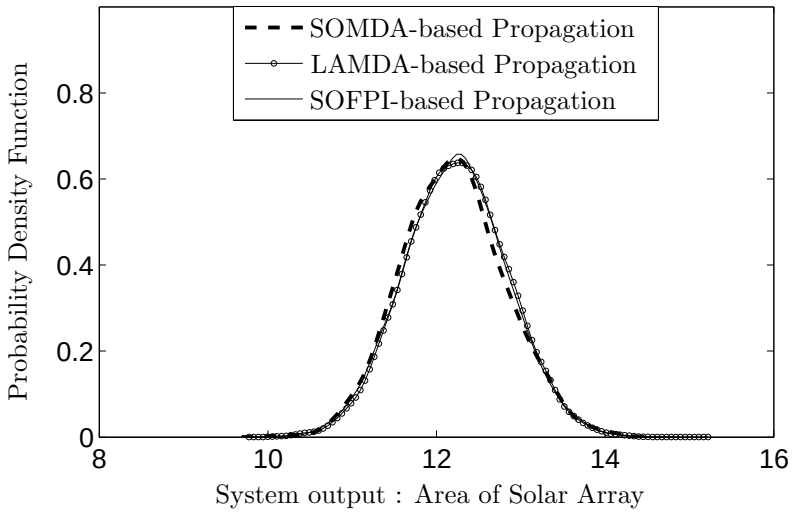


Figure 7.13: PDF of Area of Solar Array A_{sa}

decoupling the disciplinary analyses. The fully decoupled approach does not preserve one-to-one correspondence between the individual disciplinary analyses and is not suitable for further downstream analysis using the converged MDA output.

The perspective of likelihood and the ability to include input uncertainty in the construction of the likelihood function provided a computationally efficient methodology for the calculation of the PDFs of the coupling variables. The multi-disciplinary feedback analysis was reduced to a simple forward uncertainty propagation problem by replacing the feedback coupling with one-way coupling, the direction being chosen without loss of generality.

The proposed method has several advantages: (1) It provides a framework for the exact calculation of distribution of the coupling variables. (2) It retains the functional dependence between the individual disciplinary analyses, thereby utilizing the estimated PDFs of the coupling variables for uncertainty propagation, especially for downstream analyses. (3) It does not require any coupled system analysis (iterative analyses between the individual disciplines until convergence) for uncertainty propagation.

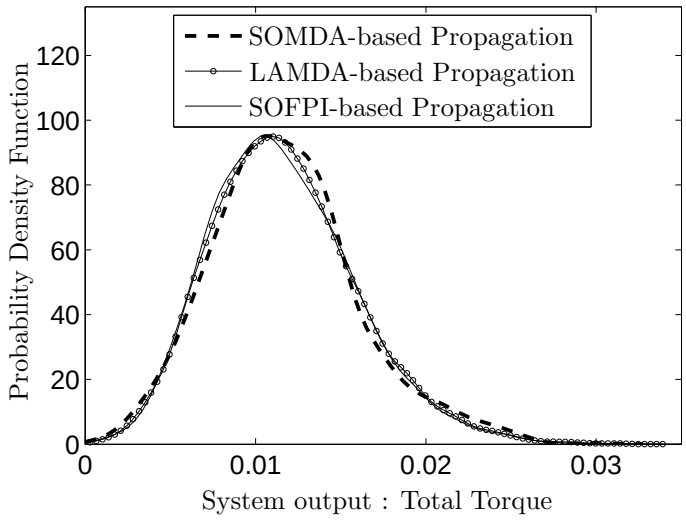


Figure 7.14: PDF of Total Torque τ_{tot}

There are several directions for future work. First, the proposed methodology was demonstrated for problems with a small number of coupling variables. The methodology is straightforward to implement when there is a vector of coupling variables as explained earlier in Section 4.2. (Recall that the fire satellite example had two coupling variables in one of the directions). However, if the coupling variable is a field-type quantity (e.g. pressures and displacements exchanged in a fluid-structure interaction problem at the interface of two disciplinary meshes), further research is needed to extend the proposed likelihood-based approach for uncertainty propagation in such multi-disciplinary problems. Second, the concept of Bayesian network can easily be extended to multi-disciplinary systems, since the feedback coupling has been replaced with equivalent unidirectional coupling. Future work needs to extend the methods developed in Chapter VI to multi-disciplinary systems with feedback coupling. Third, the likelihood-based approach can be extended to address multi-disciplinary optimization under uncertainty. Further, this chapter considered only aleatory uncertainty (natural variability) in the inputs. Fourth, future research needs

to include different types of epistemic uncertainty such as data and model uncertainty in multi-disciplinary analysis and optimization.

CHAPTER VIII

TEST RESOURCE ALLOCATION IN HIERARCHICAL SYSTEMS USING BAYESIAN NETWORKS

8.1 Introduction

In complex engineering systems, full-scale testing may not be feasible for predicting the system performance under actual operating conditions. Therefore, computational models are often used to predict the full-scale system response under actual use conditions, and thereby quantify the uncertainty associated with the system-level prediction. Research into quantification of margins and uncertainties (QMU) has the goal of enabling this overall capability [232].

Testing is essential in order to calibrate and validate the computational models used for system-level uncertainty quantification. Sometimes, it may be necessary to perform different types of tests at multiple levels of system hierarchy (component, subsystem and overall system), and for multiple types of physics (structures, heat transfer, aerodynamics, etc.) with the goal of quantifying the system performance prediction. Prior knowledge and experience can be used to subjectively describe the uncertainty in the system prediction, and this prior knowledge can be updated using testing data (i.e. experiments performed at various levels of system hierarchy and different types of physics coupling). Information in the form of testing data adds to the prior subjective knowledge, and this, in turn, reflects in the reduction of uncertainty in the system-level prediction.

Tests at different levels of the system hierarchy have different costs and different variance reduction effects. It would be ideal to select the combination of tests that lead to the maximum reduction in uncertainty, thereby minimizing the uncertainty

in the system-level prediction, while satisfying the budget constraint on testing. The objective of this chapter is to develop an analytical procedure that helps in such optimum test resource allocation. (Note that the terms "tests" and "experiments" have been used interchangeably in this chapter.)

The computational methodology for resource allocation needs to be developed immediately after model development (conceptual and mathematical) at the component, sub-system, and system levels, and it must account for the various sources of uncertainty (physical variability, data uncertainty, and model uncertainty) associated with the engineering system under study. Model uncertainty comprises of model parameter uncertainty, solution approximation errors, and model form error. While model parameter uncertainty, and solution approximation errors can be quantified and included in the test resource allocation procedure, model form error cannot be estimated before any actual tests are done. The resource allocation methodology needs to be implemented before any actual testing is performed, and hence has to use only prior knowledge and available models; model form error can be computed after the tests are conducted.

This chapter develops a Bayesian network-based methodology to efficiently solve the test resource allocation problem for multi-level, multi-disciplinary systems. The Bayesian network is an ideal choice for this purpose because it can connect multiple models, the various sources of uncertainty, and can be used for both uncertainty propagation (forward problem) and parameter calibration (inverse problem). Such an approach was earlier used by Urbina et al [233] for test resource allocation; however, Urbina et al. [233] did not calibrate multiple parameters across multiple levels of models and tests. The present chapter considers test hierarchies where each component/subsystem level of testing may be used to infer local model parameters in addition to system model parameters. Further, multi-disciplinary systems that

need to be calibrated based on tests performed at individual physics (i.e. decoupled tests) are also considered.

The key idea of the proposed methodology is to use all available component-level models and test data to quantify the uncertainty in the overall system level performance prediction. This methodology combines two types of inverse problems (model calibration and test resource optimization) and forward uncertainty propagation. The probability distributions of the model parameters are updated using the Bayesian network after collecting data through testing, and the updated distributions are propagated through the component and system models to recalculate the variance in the system performance prediction. An optimization-based procedure is then used to aid in test resource allocation by taking into consideration the reduction in variance due to testing, as well as the costs involved in testing, thereby facilitating efficient cost-benefit analysis.

The rest of this chapter is organized as follows. Section 8.2 describes the proposed methodology for test resource allocation. Sections 8.3, 8.4, and 8.5 implement the proposed methodology to a multi-disciplinary thermal vibration problem (Section 8.3), a multi-level structural dynamics problem (Section 8.4), and a multi-level, multi-disciplinary simplified space telescope mirror (Section 8.5).

8.2 Test Resource Allocation Methodology

Typically, a multi-level, multi-physics system has several parameters that influence the overall system-level output, and the uncertainty in these parameters can be updated by tests at multiple levels of the system and multiple types of physics coupling. When the posterior distributions of the parameters are propagated through the system model to calculate the overall system-level output, the posterior variance

of the overall system-level prediction can be computed. With more acquisition of data, a decreasing trend can be observed in the variance of the system-level output.

Two types of questions need to be answered: (1) What type of tests to do (which component, isolated physics, etc.)? and (2) How many repetitions of each type? Each type of test has a different testing cost and an associated reduction in the variance of system-level prediction. Further, the same type of test may need to be repeated on nominally identical specimens of the same component or subsystem. Such repetition is performed in order to account for the effect of natural variability across nominally identical specimens; while each repetition may have the same monetary cost, the associated reduction in the variance of system-level prediction may be different.

The test conducted on one subsystem is assumed to be statistically independent of another test on another subsystem; in other words, one type of test is independent of any other type. Further, for a given type of test, the repetitions across multiple replicas are also considered to be independent. It is assumed that a model is available to predict the quantity being measured in each type of test; the model may have several outputs but only that output which is measured is of concern. The overall objective is to identify how many tests of each type must be performed so as to achieve the required reduction in the variance of the system-level output. (If there are several system-level outputs, either an aggregate measure or the most critical output can be considered. Multi-objective optimization to reduce the variance of more than one system-level output needs to be considered in future work.)

8.2.1 Sensitivity Analysis

The method of sensitivity analysis has been used to quantify the sensitivity of model output to parameters. While derivative-based methods only compute local sensitivities, the method of global sensitivity analysis [36, 234, 235] can be used to

apportion the variance in the system-level output to the various sources of uncertainty, and thereby guide in the reduction of system-level prediction uncertainty.

The first step of the proposed resource allocation methodology is to use sensitivity analysis and identify those parameters that have a significant influence on the variance of the overall system-level prediction. Once the "important" parameters are identified, only those tests that aid in reducing the uncertainty in these important parameters can be performed. For example, consider a system-level output that is highly sensitive to the uncertainty in the parameters of sub-system-I but not sensitive to the parameters of sub-system-II, then it is logical to perform more sub-system-I tests than sub-system-II tests. Note that this procedure for test identification is only a preliminary approach. This approach can answer the question - "which tests to do?" In order to answer the question, "how many tests to do?", it is necessary to quantify the decrease in variance that may be caused due to a particular test. The effect of a particular test on variance reduction can be quantified by using Bayesian updating. Therefore, the proposed resource allocation methodology first uses sensitivity analysis for selection of calibration parameters and then uses Bayesian updating to quantify the effect of a test on the variance of system-level prediction.

8.2.2 Optimization Formulation

In order to solve the resource allocation problem and identify the number of tests to be performed for each type, the optimization problem can formulated in two ways, as explained below.

In the first formulation shown in Eq. 8.1, the goal is to minimize the variance of

the system-level output subject to satisfying a budget constraint.

$$\begin{aligned} &\underset{N_{test}}{\text{Minimize}} \;\; E(Var(R)) \\ &\text{s.t.} \;\; \sum_{i=1}^{q}(C_i N_i) \leq \text{Total Budget} \\ &N_{test} = [N_1, N_2 \; ... \; N_q] \end{aligned} \tag{8.1}$$

In Eq. 8.1, q refers to the number of different types of possible tests. The cost of the i^{th} ($i = 1$ to q) type of test is equal to C_i, and N_i (decision variable) denotes the number of repetitions of the i^{th} type of test. Let D_i denote all the data collected through the i^{th} type of test. Let N_{test} denote the vector of all N_i's and let $\boldsymbol{D}$ denote the entire set of data collected from all q types of tests.

Alternatively, the resource allocation problem can be formulated by minimizing the cost required to decrease the variance of the system-level output below a threshold level, as:

$$\begin{aligned} &\underset{N_{test}}{\text{Minimize}} \;\; \sum_{i=1}^{q}(C_i N_i) \\ &\text{s.t.} \;\; E(Var(R)) \leq \text{Threshold Variance} \\ &N_{test} = [N_1, N_2 \; ... \; N_q] \end{aligned} \tag{8.2}$$

In this chapter, the first formulation (Eq. 8.1) is pursued for resource allocation because the threshold level for the variance is assumed to be unknown. Using $\boldsymbol{D}$, the model parameters are calibrated and the system-level response ($R(\boldsymbol{D})$) is computed. The optimization in Eq. 8.1 calculates the optimal values of N_i, given the cost values C_i, such that the expected value of variance of the system-level prediction ($E(Var(R))$) is minimized, while the budget constraint is satisfied.

This optimization formulation uses $E(Var(R))$ as the objective function because R is a function of $\boldsymbol{D}$, which is not available before testing. Hence, random realizations

of the test data set ($\boldsymbol{D}$) are generated; each random realization is used to compute $Var(R|\boldsymbol{D})$, and the expectation over such random realizations is calculated to be the objective function, as:

$$E(Var(R)) = \int Var(R|\boldsymbol{D}) f(\boldsymbol{D}) d\boldsymbol{D} \tag{8.3}$$

where $f(\boldsymbol{D})$ is the density considered for the test data. Assuming that one type of test is performed independent of the another (i.e. a subsystem-level test is independent of a material-level test), Eq. 8.3 can be written as:

$$E(Var(R)) = \int Var(R|D_1, D_2 \ ... \ D_q) f(D_1) f(D_2) \ ... \ f(D_q) dD_1 dD_2 \ ... \ dD_q \tag{8.4}$$

where $f(D_i)$ is the density considered for the data obtained through the i^{th} test. Before any testing is done, all prior knowledge regarding the model parameters, and the mathematical models constitute the only information available for the calculation of $f(D_i)$. Therefore, $f(D_i)$ is calculated as:

$$f(D_i) = \int f(y_i|\boldsymbol{\theta}_i) f'(\boldsymbol{\theta}_i) d\boldsymbol{\theta}_i \tag{8.5}$$

where y_i represents the output of the mathematical model corresponding to the i^{th} type of test, $\boldsymbol{\theta}_i$ represents the underlying parameters, and $f'(\boldsymbol{\theta}_i)$ represents the prior knowledge regarding those parameters. Note that Eq. 8.5 is simply an uncertainty propagation problem, where the other sources of uncertainty (such as physical variability in inputs, solution approximation errors, data uncertainty) can also be included in the computation of $f(y_i|\boldsymbol{\theta}_i)$.

Eq. 8.3 – 8.5 are implemented using a numerical algorithm, where a finite number of realizations of $\boldsymbol{D}$ are generated and $E(Var(R))$ is computed over these realizations.

Then, $E(Var(R))$ can be minimized using the optimization in Eq. 8.1, and the ideal combination of tests can be identified.

Note that an inequality constraint (for the budget), and not an equality constraint, is considered in Eq. 8.1. This means that the optimal solution which minimizes $E(Var(R))$ need not necessarily exhaust the budget. Consider the simple case where there are two possible test types ($C_1 = 2$ and $C_2 = 3$), and the budget is equal to 6 cost units. There are two test combinations which exhaust the budget: (1) [$N_1 = 3$, $N_2 = 0$], and (2) [$N_1 = 0$, $N_2 = 2$]. Suppose that these two combinations lead to a value of $E(Var(R))$ which is *greater* than that achieved through the test combination [$N_1 = 1$, $N_2 = 1$]. Then, obviously the combination [$N_1 = 1$, $N_2 = 1$] must be selected because it achieves the goal of reducing $E(Var(R))$ even though it may not exhaust the budget.

8.2.3 Solution of the Optimization Problem

Eq. 8.1 is a complicated integer optimization problem, where Bayesian updating and forward propagation need to be repeated for each random realization of the test data in order to evaluate the objective function, thus increasing the computational cost several fold. In spite of the use of Gaussian process surrogate models to replace the expensive system model, high computing power is still needed to solve the optimization problem.

Integer optimization is sometimes solved using an approximation method, where the integer constraint is first relaxed, and the integers nearest to the resulting optimal solution are used in further solution of the original (un-relaxed) problem. Unfortunately, this approach is not applicable to the solution of Eq. 8.1, since the objective

function (system-level prediction variance) is defined and computed only for integer-valued decision variables (number of tests). It is meaningless to have a non-integer number of tests.

A multi-step procedure for solving the optimization problem is proposed in this chapter. Within each step, the global optimal solution is computed using an exhaustive search process, whereas across steps, a greedy algorithm is pursued. The step size is chosen in cost units, and additional steps are added until the budget constraint is satisfied.

Let the size of the first step be equal to ϕ^1 cost units; the globally optimal testing combination for this cost ($= \phi^1$) is denoted by N^1_{test}, and is calculated using exhaustive search, as:

$$\begin{aligned} &\underset{N^1_{test}}{\text{Minimize}}\ E(Var(R)) \\ &\text{s.t.}\ \ \textstyle\sum_{i=1}^{q}(C_i N_i^1) \leq \phi^1 \\ &N^1_{test} = [N_1^1, N_2^1\ ...\ N_q^1] \end{aligned} \tag{8.6}$$

The optimization procedure in the second stage is dependent on the optimal solution from the first stage, i.e. N^1_{test}. In general, the optimization for the j^{th} stage, given the solution in the previous stage (i.e. N^{j-1}_{test}), is performed for cost $= \phi^j$. Note that $\sum_j \phi^j =$ Total budget. The j^{th} optimization is formulated as:

$$\begin{aligned} &\underset{N^{j,new}_{test}}{\text{Minimize}}\ E(Var(R)) \\ &\text{s.t.}\ \ \textstyle\sum(C_i N_i^{j,new}) \leq \phi^j (i = 1 \text{ to } q) \\ &N^j_{test} = N^{j-1}_{test} + N^{j,new}_{test} \\ &N^{j,new}_{test} = [N_1^{j,new}, N_2^{j,new}\ ...\ N_q^{j,new}] \end{aligned} \tag{8.7}$$

As seen in Eq. 8.7, the decision variables for the j^{th} stage are $N^{j,new}_{test}$, i.e. those tests which need to be performed in the j^{th} stage; therefore the total number of tests is

equal to the sum of $N_{test}^{j,new}$ and N_{test}^{j-1}, i.e. the optimal number of tests in the previous stage. The same procedure is repeated until no additional test can be performed with the budget constraint satisfied.

The selection of step size for a given budget is an important issue. The true global optimal solution can be calculated by considering *one step* whose size is equal to the entire budget. However, due to the large number of possible testing combinations, this approach may be computationally infeasible. In a practical problem, several steps are considered, and the step sizes must be chosen judiciously based on (1) the costs of each type of test; (2) time required for each Bayesian update; (3) number of random realizations of data needed to compute $E(Var(R))$; and (4) the test combinations that are suitable for the chosen step size; a very small step size may not even include an expensive type of test.

8.2.4 Illustrative Example

This subsection presents a simple illustrative example, only to demonstrate the decrease of variance with testing. In order to focus on this objective, simple mathematical relationships are chosen (even the system-level response has no coupling), and measurement errors are assumed to be negligible. Other features such as coupled system response, measurements errors, solution approximation errors (while replacing the underlying physics-based model with a Gaussian process approximation), etc. are considered later in Sections 8.3 – 8.5.

The Bayesian network for this problem is exactly the same as that in Fig. 2.1. There are four independent quantities and three dependent quantities; the numerical details of this problem are specified in Table 8.1. The notation $N(\mu, \sigma)$ is used to represent a normally distributed quantity with mean μ and standard deviation σ.

Two types of tests (on two different lower levels) can be done and this information is used to update the uncertainty in the system-level response based on the tests.

Table 8.1: Numerical Details

Quantity	**Type**	**Description**
x_1	Independent	Prescribed Uncertainty : N(100,5)
x_2	Independent	To calibrate from prior : N(50, 10)
x_3	Independent	Prescribed Uncertainty : N(10,1)
x_4	Independent	To calibrate from prior : N(15, 4)
y_1	Dependent	Model prediction : $y_1 = x_1 + x_2$
y_2	Dependent	Model prediction : $y_2 = x_3 + x_4$
z	System-level response	Model prediction : $z = y_1 - y_2$
Quantity to Measure	**Cost**	**No. of Tests**
y_1	10	N_1
y_2	5	N_2

Probability distributions are assumed to be available for the inputs x_1 and x_3; if this information was not available, and only sparse and/or interval data was available for the inputs, then the likelihood-based method developed in Chapter III can be used to construct a probability distributions [236] for them. The variance of z before conducting any test (i.e. by propagating the above distributions of x_1, x_2, x_3, and x_4 through the models) is 142 units. The objective is calculate the number of tests on y_1 and y_2 (N_1 and N_2), that will lead to a minimum variance in z, subject to a total budget of \$50. Since there are only two parameters, global sensitivity analysis is not necessary, and hence, both x_2 and x_4 are chosen for calibration. The proposed optimization methodology is used for this purpose; five different stages are considered and the available budget in each stage is considered to be \$10. The results of test prioritization are given in Table 8.2 and Fig. 8.1.

At the end of the optimization procedure, the optimal combination is found to be 4 tests on y_1 and 2 tests on y_2. Further, this solution was verified by considering all other combinations (exhaustive search) of N_1 and N_2 and computing the corresponding $E(Var(R))$; for this illustrative example, this verification is numerically affordable.

Table 8.2: Resource Allocation: Results

Cumulative Cost	N_1	N_2	$E(Var(z))$
\$10	**1**	**0**	**62.0**
	0	2	127.0
\$20	2	0	53.0
	1	**2**	**46.6**
\$30	**2**	**2**	**37.6**
	1	4	46.1
\$40	**3**	**2**	**34.0**
	2	4	37.6
\$50	**4**	**2**	**32.5**
	3	4	33.8

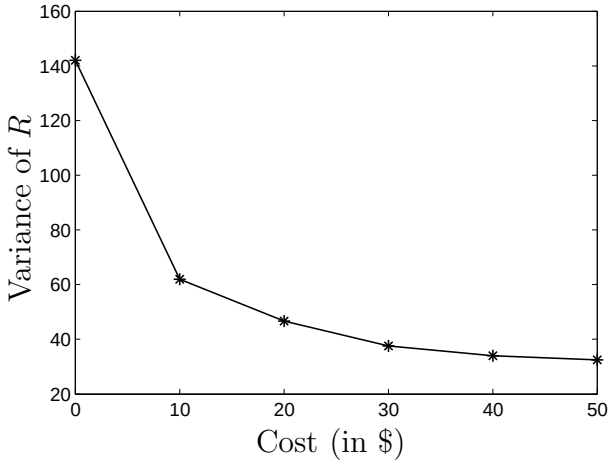


Figure 8.1: Variance vs. Cost

However, for practical examples, a few random values of $N_{test} = [N_1, N_2]$ (if not all) can be considered and it can be verified if the estimated solution is really optimal.

8.2.5 Summary of the Proposed Methodology

The various steps of the proposed methodology are summarized below:

1. **Construction of the Bayesian network:** The first step is to construct the Bayesian network that connects (1) the various component-level, subsystem-level, and system-level models; (2) the corresponding model inputs, parameters,

outputs, and solution approximation errors; and (3) the (future) experimental data and corresponding measurement errors.

2. **Sensitivity Analysis:** Global sensitivity analysis is used to identify the "important" parameters that significantly contribute to the uncertainty in the system-level response. Then, those tests which can aid in the reduction of uncertainty in these "important" parameters are selected for consideration in the optimization for test resource allocation.

3. **Bayesian updating:** The third step is to perform Bayesian updating and calibrate parameters for a particular realization of measurement data. Then, this needs to be repeated by generating multiple realizations of measurement data in order to compute the expected value of variance, as in Eq. 8.3. Since the realizations of measurement data are generated based on the model itself, model form errors are not included.

4. **Resource allocation optimization:** The third step is to perform the resource allocation optimization using the multi-step procedure developed in Sections 8.2.2 and 8.2.3. It may be useful to verify that the resultant solution is actually optimal by computing $E(Var(R))$ for few other N_{test} values.

The following sections implement the proposed test resource allocation methodology to multi-disciplinary and multi-level problems. Two different types of configurations are considered in order to emphasize on the philosophical differences involved in model development and testing of such systems.

In a multi-disciplinary system, the overall system-level is output is calculated using a multi-physics simulation, and directly indicated in the Bayesian network, since it cannot account for feedback coupling. (Even if the system-level response requires no feedback coupling, the overall system-level response can be directly indicated; in fact,

this computation is much easier than when feedback coupling is present.) The tests are always performed for individual physics without coupling. Section 8.3 discusses resource allocation for such a coupled multi-physics thermal-structural problem, representative of vibrations in solar arrays of telescopes and spacecraft booms. The tests performed for individual thermal and structural physics are used to calibrate underlying parameters, which are then used to compute the coupled system-level response.

On the other hand, in a multi-level system, the complexity of the model and underlying phenomenon increase along the hierarchy. The model used for system-level prediction is at the highest level of hierarchy and each subsequent model is at a lower hierarchy. There is a set of parameters common to the models at all levels. These parameters are calibrated using data at the lower levels (where the models and the physical phenomena are simpler relative to the system-level), and the calibrated quantities are used to predict the system-level response. For example, consider two types of tests: (1) axial test on a coupon; and (2) bending test on a beam; either/both of these tests may be used to estimate the modulus, and then predict the deflection in a thick plate, when all the three (coupon, beam, and plate) are made of the same material. Section 8.4 discusses resource allocation for such a multi-level structural dynamics problem, where tests performed on lower levels (components and subsystem) are used to calibrate the parameters, and thereby used to predict system-level response.

Section 8.5 discusses resource allocation for a multi-physics multi-level problem, where both features (tests conducted for individual physics and tests of simpler components or conditions) are used for the calibration of parameters and prediction of the system-level response.

8.3 Multi-disciplinary System

8.3.1 Description of the Problem

This coupled-physics thermal vibration example illustrates a laboratory experiment which can be used to study and simulate the behavior in solar arrays of telescopes and spacecraft booms [208]. In this experiment, a thin walled circular tube is rigidly attached at its top and supports a concentrated mass at its bottom. The tube and the mass are initially at rest and a constant heat flux is applied on one side along the length of the tube. The application of the heat flux causes an increase in the temperature on the incident surface while the unheated side remains at the initial temperature. The temperature gradient causes the beam to bend away from the lamp, due to the thermal moment. The displacement of the beam, in turn, changes the distribution of temperature along the length of the beam, leading to a change the temperature gradient and the thermal moment, which in turn affects the flexural behavior. Thus the combination of heat transfer and flexural mechanics leads to oscillations of the beam. The set up of this experiment is shown in Fig. 8.2.

The temperature at the tip mass (T_m) is given by the following differential equation:

$$\frac{\partial T_m}{\partial t} + \frac{T_m}{\tau} = \frac{T^*}{\tau}(1 - \frac{v(x,t)}{\beta^*}) \tag{8.8}$$

In Eq. 8.8, $v(x,t)$ represents the displacement of the beam as a function of length and time. Thornton [208] explains how to calculate the parameters T^*, τ, β^* as a function of the incident solar flux (S).

The displacement $v(x,t)$ can be related to the displacement of the tip mass $V(t)$ as:

$$v(x,t) = (\frac{3x^2}{2l^2} - \frac{x^3}{2l^3})V(t) \tag{8.9}$$

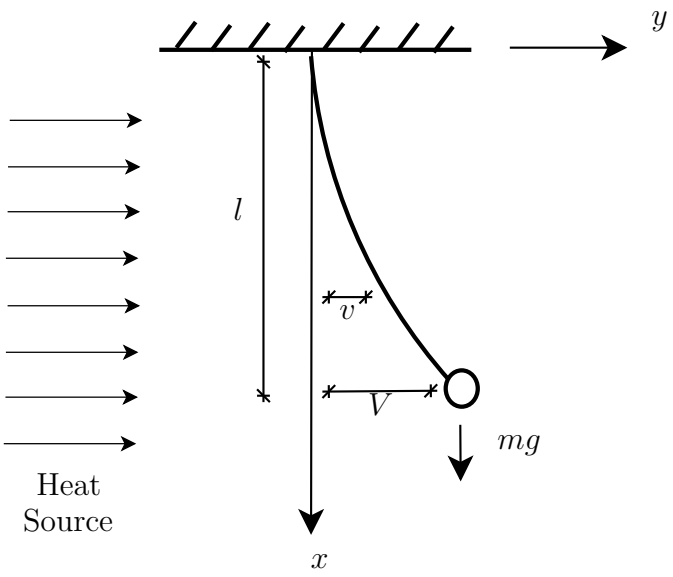


Figure 8.2: Thermally Induced Vibration

The tip mass displacement $V(t)$, in turn, depends on the forcing function as follows:

$$\ddot{V} + 2\xi\omega_0\dot{V} + (\omega_0^2 + \frac{6g}{5l})V = \frac{F(t)}{m} \tag{8.10}$$

In Eq. 5, ξ is the damping ratio, and ω_0 is the angular frequency. The forcing function $F(t)$ depends on the thermal moment which in turn depends on the temperature, thereby casing coupling between the thermal equation and the structural equation. These relations are shown in the following equations:

$$F(t) = -\frac{3}{l^3}\int_0^l \int_0^x M(u,t)dudx \tag{8.11}$$

$$M(x,t) = \int E\alpha T_m(x,t)cos(\Phi)ydA \tag{8.12}$$

In Eq. 8.12, E is the elastic modulus, α is the coefficient of thermal expansion, Φ is the angle of incident flux on the cross section, y is the distance from the center

Table 8.3: Calibration Quantities: Thermal Vibration Problem

Symbol	**Quantity**	Property	**Prior CoV**
E	Elastic modulus	Structural	0.1
c	Independent	Thermal	0.1
ξ	Independent	Structural	0.1
r	Independent	Geometric	0.03
e	Dependent	Thermal	0.1

of the cross section and the integral is over the area of the cross section A. Refer Thornton [208] for a detailed description of this problem.

The overall objective of test resource allocation is to minimize the variance of the system-level output (R), which is defined to be the ratio of displacement amplitudes at two different time instants for the coupled system when the incident solar flux (S) is 2000 W/m^2. If $R < 1$, the system is stable with oscillations diminishing as a function of time. If $R > 1$, the system is unstable, commonly referred to as flutter, an undesirable scenario. While Gaussian process surrogate model is constructed to calculated R, individual physics predictions are performed using the above physics-based models.

There are several parameters (both thermal and structural) in the above equations, that can be calibrated using test data. The method of sensitivity analysis is used to identify five parameters, which significantly contribute to the uncertainty in the system-level prediction. The prior means are based on [208], and the assumed coefficients of variation (CoV) are tabulated in Table 8.3; note that the radius being a geometric property has a lower CoV. The calibrated parameters are then used to quantify the uncertainty in R.

The calibration parameters need to be estimated during test data; four different types of tests are considered, as shown in Table 8.4. The total budget available for

Table 8.4: Types of Tests: Thermal Vibration Problem

Test type	Physics	Calibrate	Input - Output	Cost	No. of tests
Material-level	Thermal	c	Heat-Temperature rise	\$100	N_{m_1}
Material-level	Structural	ξ	Amplitude decay	\$100	N_{m_2}
Subsystem-level	Thermal	c, e, r	Heat-Temperature rise	\$500	N_T
Subsystem-level	Structural	c, e, r	Acceleration	\$500	N_F

testing is assumed to be \$2000. It is assumed that the entire multi-disciplinary system cannot be tested.

The calibration quantities, the model predictions, and the test data are connected through a Bayesian network, as shown in Fig. 8.3.

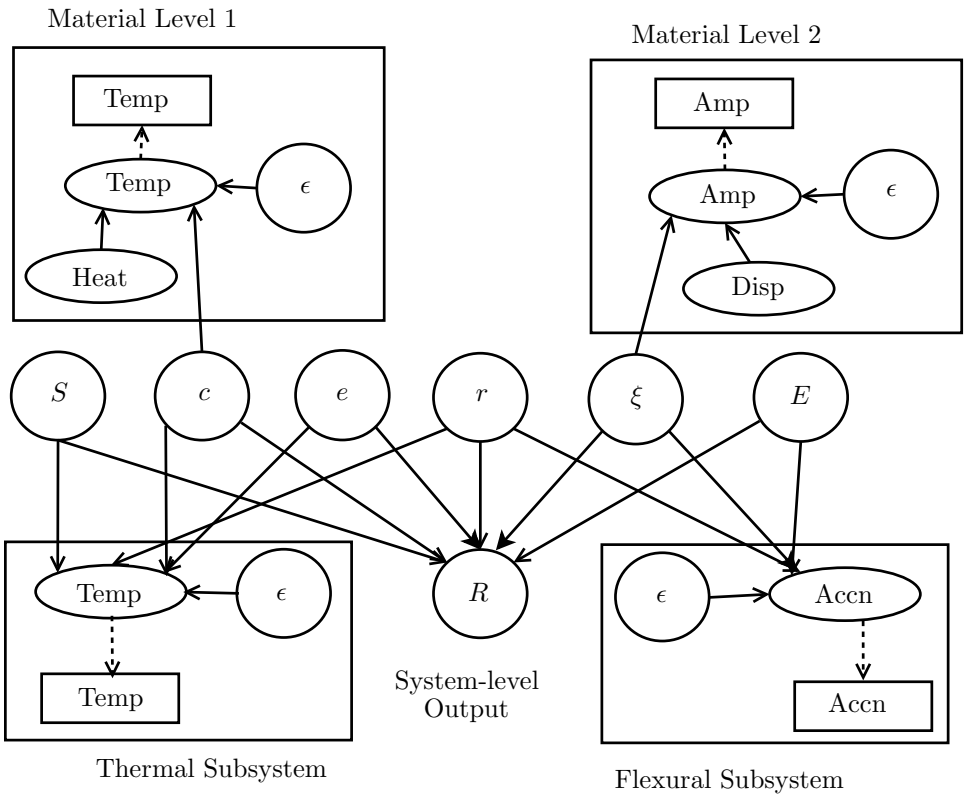


Figure 8.3: Thermal Vibration: Bayesian Network

In the Bayesian network in Fig. 8.3, "Temp" refers to temperature, "Accn" refers to the acceleration, "Disp" refers to the displacement, and "Amp" refers to the amplitude of vibration. Measurement errors (ϵ) are assumed to have a standard deviation

that is equal to ten percent of the model prediction. This Bayesian network is used for uncertainty quantification, Bayesian updating and resource allocation.

8.3.2 Resource Allocation

The objective is to calculate the number of tests that lead to maximum reduction in variance in R. Let N_{test} denote the number of tests, where $N_{test} = [N_{m_1}, N_{m_2}, N_F, N_T]$; where N_{m_1} is the number of material level temperature tests, N_{m_2} is the number of material level pluck tests, N_F is the number of flexural subsystem tests, and N_T is the number of thermal subsystem tests. Let $D = [D_{m_1}, D_{m_2}, D_F, D_T]$ denote the test measurements. The optimization problem for resource allocation can be formulated as shown in Eq. 8.13

$$\begin{aligned} &\underset{N_{test}}{\text{Minimize}}\ E(Var(R)) \\ &\text{s.t. } 100(N_{m_1} + N_{m_2}) + 500(N_F + N_T) \leq 2000 \\ &N_{test} = [N_{m_1}, N_{m_2}, N_F, N_T] \end{aligned} \tag{8.13}$$

The above optimization is solved using the multi-stage optimization procedure discussed in Sections 8.2.2 and 8.2.3. Four stages and a budget of \$500 for each stage are considered, thereby accounting for the total budget of \$2000. Each stage has 8 options (as against two in the mathematical example in Section 8.2); only the optimal solution in each stage is shown.

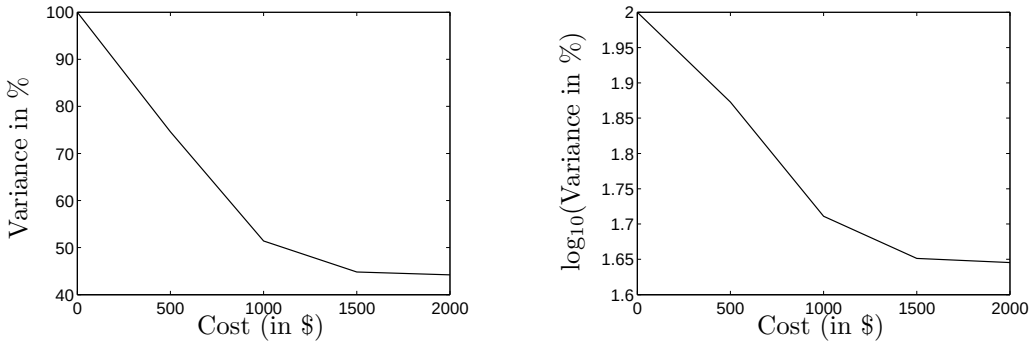
Note that Table 8.5 expresses the expectation of variance of R in terms of percentage of the variance before any testing; this variance is equal to 5.69 x 10^{-7}; since R is a ratio, this variance is dimensionless.

For a \$2000 budget, it is seen that one temperature test, nine pluck tests, one thermal subsystem test and one flexural subsystem test are required to achieve the maximum reduction in the variance of R. The results show that while it is useful to

Table 8.5: Resource Allocation Results: Thermal Vibration Problem

Stage No.	N_{m_1}	N_{m_2}	N_F	N_T	$E(Var(R))$ (in %)
No tests	0	0	0	0	100.0
Stage 1 : \$500	1	4	0	0	74.6
Stage 2 : \$1000	1	4	1	0	51.4
Stage 3 : \$1500	1	4	1	1	44.8
Stage 4 : \$2000	1	9	1	1	44.2

do all the tests, repeating the pluck test which calibrates structural damping, is not only cheap but also leads to effective decrease in the variance of R. The decrease of variance with cost is shown in Fig. 8.4.



(a) Variance with cost (b) Log(Variance) with cost

Figure 8.4: Decrease of Variance with Cost

It is seen that the reduction in variance using the last \$1000 (i.e. from \$1000 to \$2000) was much smaller when compared to the reduction in variance using the initial \$1000. Such information is useful for budgeting purposes, since all the above computation (and practical resource allocation) is done before any test is actually conducted.

8.4 Multi-level System

8.4.1 Description of the Problem

A three-level structural dynamics developed at Sandia National Laboratories [200] is considered as shown in Fig. 8.5.

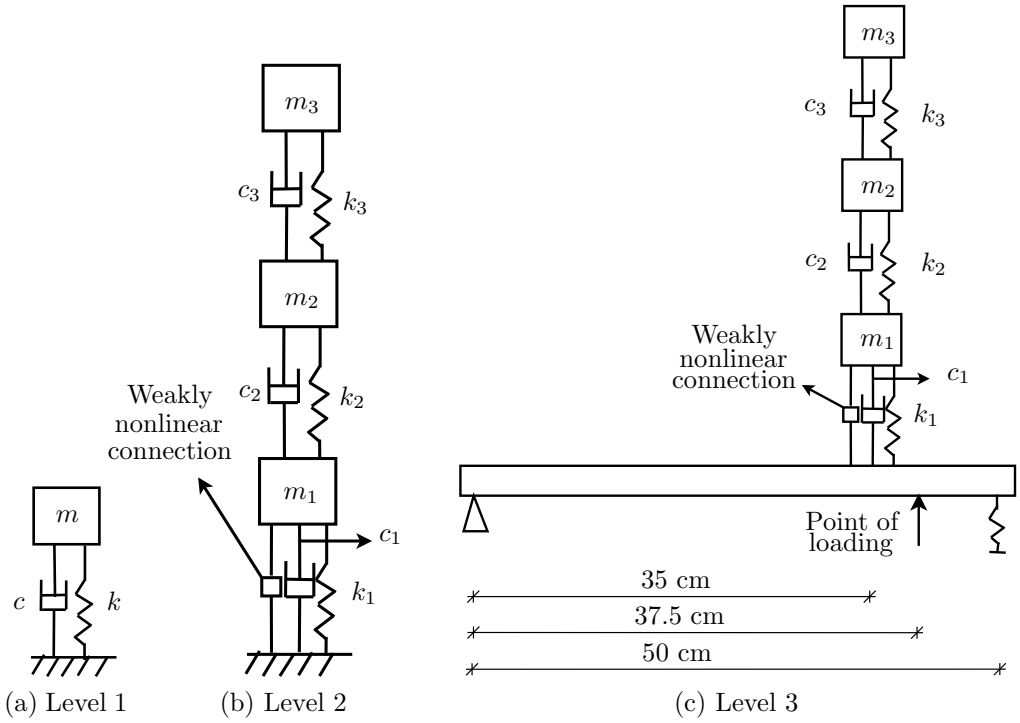


Figure 8.5: Multi-level Structural Dynamics Problem

The first level (component) consists of a single spring-mass-damper. Three such spring-mass dampers are integrated to form a spring-mass-damper subsystem in the second-level. In the third level, the integrated spring-mass-damper subsystem is mounted on a beam to form the overall system.

The models to represent the first two levels are straightforward [204]. Red-Horse and Paez [200] describe in detail the modeling and simulation of the overall system (third-level). The overall objective is test resource allocation to minimize the variance of the system level output (R) which is defined to be the maximum acceleration

of mass m_3, when a random force is applied as specified in [200]. The first-level and second-level responses are computed using physics-based models while the third-level and system-level responses are computed by constructing two Gaussian process surrogate models.

In this numerical example, the stiffness values of the three masses, i.e. k_1, k_2, and k_3 are all the parameters that need to be calibrated with test data; since all parameters are calibrated, sensitivity analysis is not used in this example. The numerical values (in SI units) of three calibration parameters are summarized in Table 8.6.

Table 8.6: Model Parameters: Structural Dynamics Problem

Number	Mass (m) (in kg)	Damping (c) (in Ns/m)	Prior Mean of Stiffness (μ_k) (in N/m)	Prior Std. Dev. of Mean (σ_k) (in N/m)
1	0.0125	0.023	5600	560
2	0.0193	0.021	11000	1100
3	0.0351	0.031	93000	9300

The mass of the beam is taken to be 0.1295. Further numerical details of the beam are given in [200].

Data for calibration is assumed to be available through five different types of tests. The details of these different types of tests are provided in Table 8.7. For each test, a sinusoidal load (amplitude=10000 and angular velocity = 10 rad s^{-1}) is used. For the first and second level tests, the sinusoidal load is applied at the base; for the third level test, the sinusoidal load is applied as specified in [200].

Table 8.7: Types of Tests: Structural Dynamics Problem

Test Type	Description	Model prediction	Data	Cost	No. Tests
Level-1	Only mass m_1	acceleration (x_{11})	D_{11}	\$100	N_{m_1}
Level-1	Only mass m_1	acceleration (x_{12})	D_{12}	\$100	N_{m_2}
Level-1	Only mass m_1	acceleration (x_{13})	D_{13}	\$100	N_{m_3}
Level-2	3-mass assembly	acceleration of m_3 (x_2)	D_2	\$500	N_2
Level-3	3-mass assembly on beam	acceleration of m_3 (x_3)	D_3	\$1000	N_3

The model predictions, experimental data, and the calibration quantities are connected using the Bayesian network, shown in Fig. 8.6. The corresponding experimental errors are denoted by ϵ_{11}, ϵ_{12}, ϵ_{13}, ϵ_2, and ϵ_3 respectively, and assumed to be equal to ten percent of the prediction.

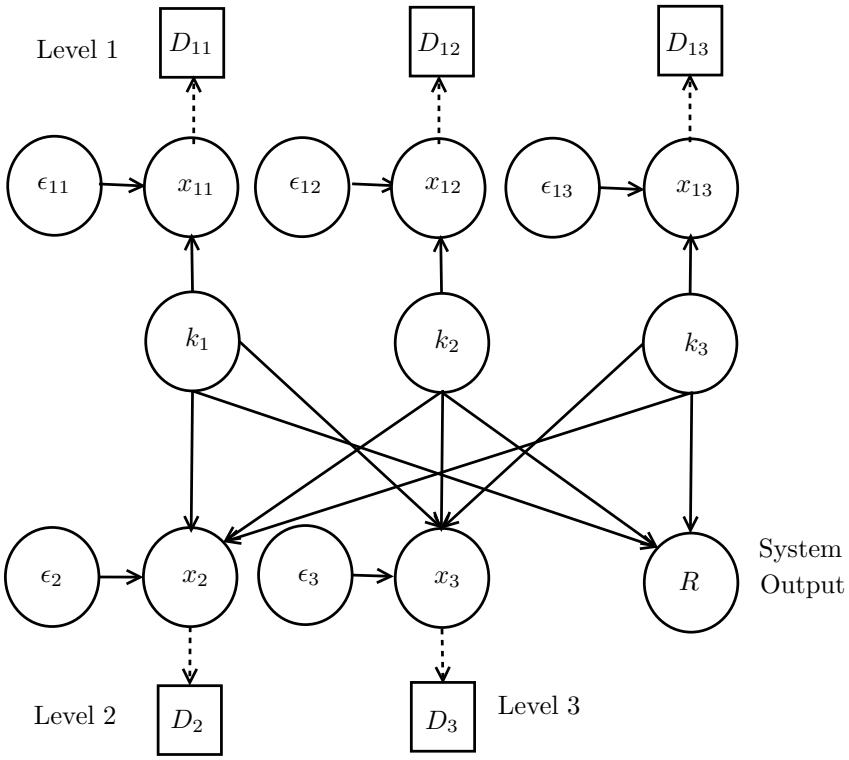


Figure 8.6: Bayesian Network: Structural Dynamics Problem

This Bayesian network can be used for uncertainty quantification, Bayesian updating, and resource allocation, as explained below.

8.4.2 Resource Allocation

In the resource allocation problem, testing is yet to be done and hence realizations of future experimental data are generated randomly. Then, $E(Var(R))$ is computed so as the identify which set of tests will lead to the maximum reduction in variance. Let $N_{test} = [N_{m_1}, N_{m_2}, N_{m_3}, N_2, N_3]$. The optimization problem for resource allocation can be formulated as shown in Eq. 8.14.

Table 8.8: Resource Allocation Results: Structural Dynamics Problem

N_{m_1}	N_{m_2}	N_{m_3}	N_2	N_3
0	1	4	1	0
0	4	1	1	0
0	3	2	1	0
0	2	3	1	0
1	1	3	1	0
3	1	1	1	0
1	3	1	1	0
1	2	2	1	0
2	1	2	1	0
2	2	1	1	0

$$\begin{aligned} &\underset{N_{test}}{\text{Minimize}}\ E(Var(R)) \\ &\text{s.t. } 100(N_{m_1} + N_{m_2} + N_{m_3}) + 500N_2 + 1000N_3 \leq 1000 \\ &N_{test} = [N_{m_1}, N_{m_2}, N_{m_3}, N_2, N_3] \end{aligned} \tag{8.14}$$

First, the resource allocation is solved for a budget of \$1000. There are 54 possible testing combinations and out of these 54, ten testing combinations lead to the same minimum variance of system-level output R, approximately 0.8% of the variance before testing. These combinations are given in Table 8.8. The value of $E(Var(R))$ for these ten cases are close enough that it is not possible to determine whether the difference is due to reality or due to sampling/numerical errors.

It is a subjective decision as to which one of these ten test combinations is selected. However, all ten combinations unanimously suggest that no tests are needed for the overall system and one test is needed for the second level three spring-mass-damper subsystem. The first four rows in Table 8.8 suggest that testing is not needed for the first spring-mass-damper. However, it may be desirable to have at least one test for each component, and hence one amongst the latter six options may be preferred.

It was also found that an extra budget of \$1000 caused no further reduction in the

variance of R. If the available budget is $ 2000, a subjective decision may be made to conduct the full system test (which costs $1000) in order to further improve the confidence in uncertainty quantification.

8.5 Multi-level, Multi-disciplinary System

8.5.1 Description of the Problem

A simplified space telescope mirror problem is considered as an example of a multi-level, multi-disciplinary system. As shown in Figure 8.7, it consists of three components - leg, mirror and plate - which are integrated to form the overall system, which can also be decomposed into various pieces as shown in Figure 8.7.

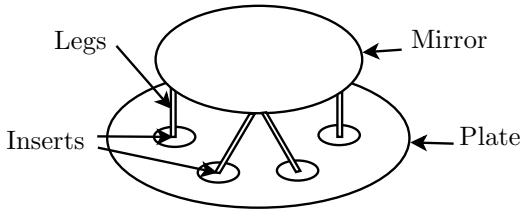


Figure 8.7: Simplified Space Telescope Mirror Problem

The system is tested under two types of physics – mechanical (due to gravity loading) and thermal (due to solar flux), that interact with each other, and affect the optical performance of the mirror. Eight different types of tests are considered, as tabulated in Table 8.9. It is assumed that the full system test under combined mechanical and thermal loading cannot be performed.

The system was simulated using the Sierra multi-physics mechanics simulation suite developed at Sandia National Laboratories [237]. The thermal and the structural properties of the system affect the overall optical performance of the telescope mirror. The thermal and structural meshes were independent, with different, and programmable mesh densities. In each case, as appropriate, the Sierra code Aria was

Table 8.9: Types of Tests: Telescope Mirror

Test No.	Component	Physics	Cost Units
1	Only leg	Gravity sag	\$1
2	Only mirror	Gravity sag	\$3
3	Only plate	Gravity sag	\$1
4	Entire assembly	Gravity sag	\$15
5	Only leg	Solar flux	\$10
6	Only mirror	Solar flux	\$30
7	Only plate	Solar flux	\$10
8	Entire assembly	Solar flux	\$150

used for thermodynamics, heat transfer, and radiation modeling, and the Sierra code Adagio was used for solid mechanics and quasi-static transient dynamics. For the purpose of this study, the optical system output was simply taken to be the deformation of the mirror at the center of the mirror. For this study, each Sierra simulation was wrapped within a DAKOTA [238] script to generate input-output data, that were later used to build Gaussian process surrogate models. Nine different surrogate models are constructed; eight of them to make predictions corresponding to the tests in Table 8.9, and the ninth is for system-level prediction. The overall system output (denoted by R) is chosen to be the deflection of mirror relative to the center of the plate under both gravity and solar flux; this is equal to the sum of individual deflections under gravity sag and solar flux. Deflection is here a proxy for performance metrics such as wavefront error, focus drift, or other system-level characteristics that cannot be well-represented without coupling the structural, thermal, and optical models. The test data and the Gaussian process models can be connected through a Bayesian network, as shown in Fig 8.8.

Consider the Bayesian network in Fig. 8.8. The quantities θ_L, θ_M, θ_P and x_L, x_M, x_P refer to the model parameters and inputs of the leg, mirror, and plate components respectively; note that each of these quantities is a vector since a component may have more than one parameter/input. Each model parameter vector consists of the

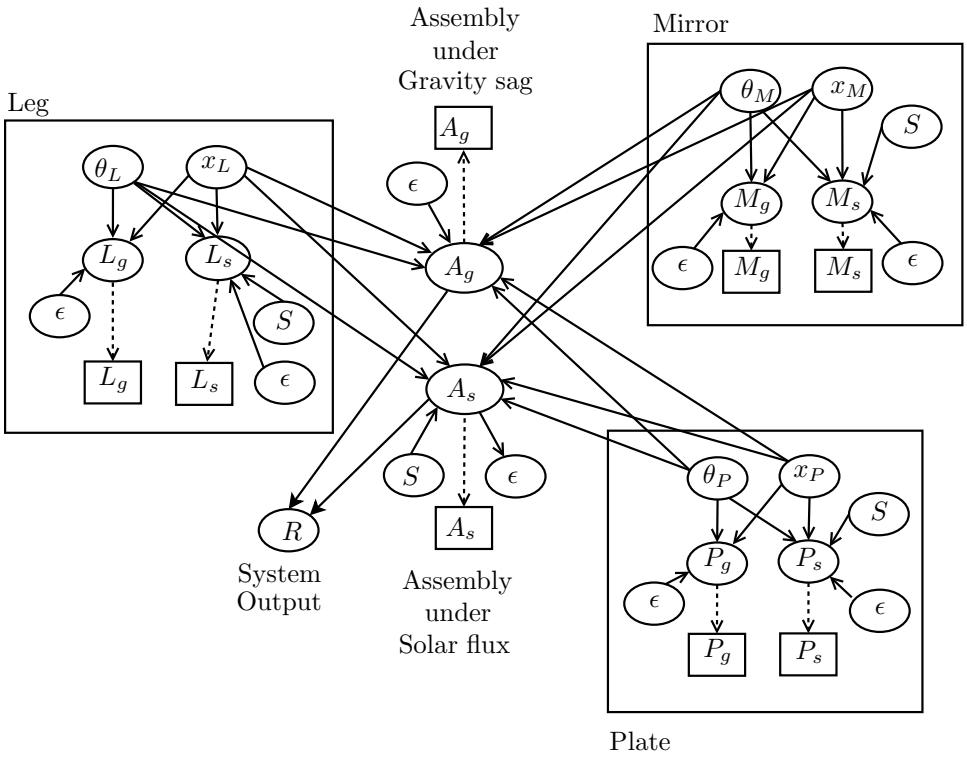


Figure 8.8: Telescope Mirror Problem: Bayesian Network

corresponding component's density, Young's modulus, Poisson's ratio, coefficient of thermal expansion, thermal conductivity, specific heat, and emissivity. Thus, this example demonstrates the scalability of the proposed methodology by considering eight possible types of tests and twenty-one different parameters. The model predictions (for leg, mirror, plate, and assembly) are denoted by L_g, M_g, P_g, A_g (mechanical loading due to gravity sag) and L_s, M_s, P_s, A_s (thermal loading due to solar flux) respectively. Though the same symbol ϵ has been used to denote the difference between model prediction and observation throughout the Bayesian network, the statistics of ϵ is different for different tests, and equal to ten percent of the model prediction.

8.5.2 Sensitivity Analysis

Sensitivity analysis is particularly important in this example because there are possible possible calibration parameters. The sensitivities of the system-level output R to all the parameters – previously mentioned twenty-one model parameters (θ_L, θ_M, θ_P) – are quantified using sensitivity analysis based on DAKOTA [238]. The "important" parameters based on the results of sensitivity analysis are tabulated in Table 8.10.

Table 8.10: Sensitivity Analysis for Coupled System: Gravity Sag and Solar Flux

Model Parameter	**Importance Measure**	**Cumulative Importance**
Mirror Young's Modulus	0.8021	0.8021
Leg Emissivity	0.0277	0.8298
Mirror Poisson's Ratio	0.0235	0.8533
Mirror Density	0.0112	0.8645

Further, the input solar flux has an importance measure of 0.12; however this is not a model parameter. From Table 8.10, it can be seen that four model parameters, along with the solar flux, account for more than 98% of the variance of the system output. These four parameters, i.e. mirror elastic modulus, leg emissivity, mirror Poisson's ratio, and mirror density, are chosen to be calibrated through testing.

8.5.3 Test Resource Allocation

The goal of the resource allocation problem is to select tests that minimize the variance of the overall system output (R) under both gravity sag and solar flux. There are four types of tests that can be useful to calibrate the aforementioned "important" parameters; these tests are the gravity sag assembly level test (number of tests = N_{GSA} and each test costs 15 units), the solar flux assembly level test (number of tests = N_{SFA} and each costs 150 units), the solar flux leg component test (number of tests

$= N_{SFL}$ and each costs 10 units) and the gravity sag mirror component test (number of tests $= N_{GSM}$ and each costs 3 units). In each test, the displacement of the mirror under the given loading is measured; correspondingly four different Gaussian process surrogate models are constructed to obtain model predictions. Also, a total budget of 150 cost units is assumed to be available. The optimization for test resource allocation is written as:

$$\begin{aligned} &\underset{N_{test}}{\text{Minimize}}\ E(Var(R)) \\ &\text{s.t. } 15N_{GSA} + 150N_{SFA} + 10N_{SFL} + 3N_{GSM} \leq 150 \\ &N_{test} = [N_{GSA}, N_{SFA}, N_{SFL}, N_{GSM}] \end{aligned} \tag{8.15}$$

The results of test resource allocation are given in Table 8.11. Similar to the previous sections, this is a multi-stage optimization. In each stage, 30 cost units are considered, and there are seven options to exhaust a budget of 30 cost units in each stage, and the optimal solution in each stage is shown. Table 8.11 presents the variance in terms of percentage of the variance of R before testing (which was equal to 1.33 x 10^{-12} m^2). N_{test} is the vector of (N_{GSA}, N_{SFA}, N_{SFL}, N_{GSM}).

Table 8.11: Multi-stage Optimization

N_{test}	Cost Units	$E(Var(R))$ (in %)
(0, 0, 1, 6)	28	12.3
(1, 0, 1, 11)	58	8.7
(3, 0, 1, 11)	88	7.4
(4, 0, 1, 16)	118	6.6
(5, 0, 1, 21)	148	6.1

The results of the test resource allocation optimization recommend 5 assembly–level tests under gravity sag, 1 leg component test under solar flux, and 21 mirror component tests under gravity sag in order to minimize the system output variance.

The decrease in variance with cost, based on the optimal solution in each stage, is shown in Fig. 8.9.

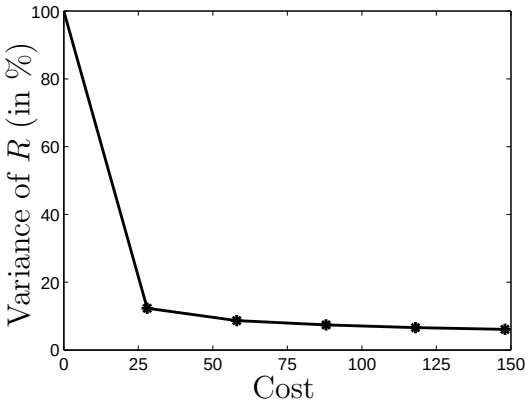


Figure 8.9: Variance vs. Cost

If the assembly–level tests (one under gravity sag and one under solar flux) were alone performed, then the variance decreases to 33.6% of the original value, at a cost of 165 units. Hence, it is evident that the proposed methodology achieves better performance (higher reduction in variance) at a lower cost.

It is seen from Fig. 8.9 that there is little improvement in the system variance after testing worth 58 cost units. At that point, the results recommend performing 1 assembly–level test under gravity sag, 1 leg component test under solar flux and 11 mirror component tests under gravity sag. Hence, subsequent tests do not significantly aid in the reduction of variance in R. If the alternate optimization formulation (Eq. 8.2) with a threshold variance lower than 5% of the prior variance had been chosen for resource allocation, then it may have been impossible to satisfy the variance constraint. Therefore, the optimization formulation in Eq. 8.1 may be preferred, since it provides an estimate of the variance with cost.

8.6 Conclusion

Testing at the component, subsystem and system levels is important in the context of uncertainty quantification in multi-level systems. When the systems are multi-disciplinary, it is important to conduct tests for both individual and combined physics. But rarely is it feasible to conduct every imaginable test, either due to schedule or cost considerations. This chapter developed an analytical procedure to aid in deciding which tests to conduct, especially for complex and expensive systems. A Bayesian network is used to connect multiple models and test data at different levels, and also include the various sources of error and uncertainty. The steps of the proposed methodology can be summarized as follows: (1) connect models and experiments at multiple levels efficiently through a Bayesian network; (2) systematically account for and include natural variability, data uncertainty, and solution approximation errors; (3) predict the overall uncertainty in the system model prediction; and (4) optimize resource allocation for test selection and identify the most effective tests to reduce overall system model uncertainty.

A lower level test can easily isolate individual components and hence, the model parameters can be effectively updated, leading to a significant reduction in the variance of the system-level prediction. However, such a test would not account for interactions between higher level models and the corresponding parameters. In contrast, a higher level test would include the effects of interaction between multiple subsystem-level and component-level models. However, the calibration of parameters across multiple models may be difficult and may not lead to a significant reduction in the variance of the system-level prediction. The proposed test resource allocation procedure trades off between lower level tests and higher level tests by accounting not only for the resultant reduction in variance of the system-level prediction but also the testing costs.

Future work needs to address three major issues. The first deals with computational effort. As the number of calibration variables increases and the number of types of tests increases, the numerical difficulties involved in the numerical solution of the optimization problem increase; efficient numerical methods need to be developed for this purpose. The second deals with test design; having identified the number of tests, the next step would be to design them in order to maximize the uncertainty reduction in the system level prediction. The entropy-based approach developed by Jiang and Mahadevan [239] for the design of validation experiments may be investigated for this purpose. Finally, though the three numerical examples presented in this chapter considered different features (multiple levels of complexity and coupled physics interactions) representative of practical applications, it is necessary to further investigate the extension of the proposed methodology for realistic engineering problems.

An important observation in this chapter was that different parameters have different sensitivities to different types of tests. Further, it was shown in Chapter IV, that the presence of other sources of uncertainty also affect the parameter uncertainty. If the goal is to reduce the uncertainty in the system-level response by reducing the uncertainty in the system parameters, it is not only sufficient to study the sensitivity of the system-output to the parameters, but also necessary to study the sensitivity of calibration parameters to the other sources of uncertainty and the calibration data itself. While the former is addressed using the global sensitivity analysis method discussed in Section 2.6, the latter has not been studied at all. This dissertation refers to this topic as "inverse sensitivity analysis" and studies it in detail in the following chapter.

CHAPTER IX

INVERSE SENSITIVITY ANALYSIS

9.1 Introduction

The topic of parameter estimation, i.e. inferring an unobservable (or difficult to measure) quantity through the measurement of a dependent variable, has been been a significant topic of interest over several years [12]. Researchers have pursued two class of methods for parameter estimation. The first class of methods is purely based on explicit functional inversion of the mathematical model (i.e. forward model) used to represent the dependent variable as a function of the parameters to be estimated. This approach becomes infeasible with increasing non-linearity of the aforementioned mathematical model, and with the presence of additional sources of uncertainty [122]. In fact, this is the case with practical engineering systems whose response not only requires the solution of complicated differential equations and but is also stochastic due to the presence of several other sources of uncertainty. Hence, it becomes impossible to directly invert the system response function in order to estimate the system parameters. Hence, it becomes necessary to the second class of methods which are based on principles of statistics; these methods do not use direct inversion and estimate the parameters simply using repeated evaluations of the forward model.

Statistical methods for parameter estimation were discussed in detail in Section 4.3 in Chapter IV. Since Chapter IV, a Bayesian approach for model calibration and parameter estimation has been pursued in this dissertation because it can rigorously account for the various sources of uncertainty – physical variability, data uncertainty, and model uncertainty – in a systematic manner, and subjectively estimate the entire PDF of the model parameters, thereby quantifying the uncertainty in them.

Quantifying the uncertainty in the parameter is not equivalent to simply computing the probability distribution of the parameter. Uncertainty quantification must also analyze the various sources of uncertainty that lead to the uncertainty in the parameter and thereby, compute the individual contributions of the various sources of uncertainty to the uncertainty in the parameter. This topic of apportioning uncertainty has received considerable treatment in the case of forward problems, i.e. apportioning the uncertainty in the model output to the uncertainty in the inputs. In the case of forward problems, while a preliminary approach would be to analyze the sensitivity of the variance of the model output by suppressing the uncertainty in each of the inputs, researchers have pursued a rigorous variance-based global sensitivity analysis [36] approach for this purpose. This approach was briefly reviewed in Section 2.6.

In the case of inverse problems, researchers have realized the importance of model calibration under uncertainty; while there are several published journal articles that deal with quantifying the uncertainty in the model parameters, the aspect of sensitivity analysis of model parameters has not received much attention. The focus of this chapter is to develop a statistical methodology to analyze and apportion the uncertainty in model parameters to the other sources of uncertainty. This objective is significantly different from "forward sensitivity analysis", and hence the issue of sensitivity analysis of model parameters is referred to as "inverse sensitivity analysis" in this chapter. Note that both of these are approaches are "global" because they account for the entire distribution of the uncertain quantities.

Consider the computational model $Y = G(\boldsymbol{X}; \boldsymbol{\theta})$; though this model is deterministic (the output y is single-valued for a given realization of inputs $\boldsymbol{X}$ and parameters $\boldsymbol{\theta}$), the inputs $\boldsymbol{X}$ are uncertain with probability density function (PDF) given by $f_{\boldsymbol{X}}(\boldsymbol{x})$, and the parameters $\boldsymbol{\theta}$ are unknown. Prior PDFs ($f'(\boldsymbol{\theta})$) are assumed for the

model parameters, data D is collected on the output, and Bayesian inference is used to calculate the posterior PDF ($f''(\boldsymbol{\theta})$). The uncertainty in the inputs ($f_{\boldsymbol{X}}(\boldsymbol{x})$) can be included in a systematic manner during this calibration procedure. The details of this calibration were discussed in Section 4.3.

The posterior PDFs of the model parameters depend on (1) the uncertainty in the inputs; (2) the data available for parameter estimation; and (3) the choice of prior of model parameters. For a given model parameter (denoted by θ^i), the effect of the choice of its prior is easy to quantify by computing the posterior corresponding to every choice of a prior. Further, the prior simply has a multiplicative effect during the computation of the posterior, and hence, its effect can be studied easily. On the other hand, the effect of uncertainties and data on the posterior cannot be easily quantified. This chapter focuses on these two tasks. It is important to note that that the posterior of a particular model parameter (θ^i) depends only on the prior of the other model parameters (denoted by $\boldsymbol{\theta}^{-i}$) and not on their posterior. As a result the posterior variance of θ^i depends only on the prior of $\boldsymbol{\theta}^{-i}$; hence with respect to θ^i, $\boldsymbol{\theta}^{-i}$ can be just treated similar to the inputs $\boldsymbol{X}$, i.e. just as an additional uncertainty described by the prior PDF.

The sensitivity analysis of a particular model parameter θ^i, has multiple facets:

1. Sources of uncertainty (prescribed uncertainty for inputs x and prior uncertainty for other model parameters $\boldsymbol{\theta}^{-i}$)

2. Data for calibration ((i) number of data; and (ii) numerical values of data)

The former, i.e. analyzing the sensitivity of model parameter with respect to the other sources of uncertainties, is similar to the global sensitivity analysis methodology, the conceptual difference being that forward GSA apportions the uncertainty in the model output to the various sources of uncertainty, whereas the goal of the

present chapter is to apportion the uncertainty in the estimated model parameter to the various sources of uncertainty. Such an analysis is useful in several ways. First, It helps to identify important contributors of uncertainty. For example, there may be some sources of epistemic (reducible) uncertainty which contribute to the uncertainty in the parameter and the amount of contribution can be quantified. Second, if a particular source of uncertainty has insignificant sensitivity, then that source of uncertainty can be treated to be deterministic; this in turn reduces the computational effort by reducing the dimensionality. Third, if a design of experiments need to be performed over the space of the uncertain quantities, then the results of inverse sensitivity analysis can be used; if a particular quantity is not significant, then it may not be necessary to select multiple realizations of that quantity and vice-versa.

The benefits of inverse GSA are evident, and this topic is yet to be addressed in the literature. The methods of forward sensitivity analysis are alone well-established in the literature, and these cannot be directly applied to inverse sensitivity analysis. This issue is discussed later in Section 9.2, where it is demonstrated that the equations for variance decomposition are not directly applicable for inverse sensitivity analysis. The first contribution of this chapter is to develop a statistical methodology for such sensitivity analysis.

The latter, i.e. analyzing the sensitivity of model parameter with respect to the available data, is important from the perspective of variance reduction, which was the focus of Chapter VIII. Ideally (if the model is identifiable), it is intuitive that an increase in the number of calibration data points will lead to a decrease in the variance of the model parameter. For the purpose of test planning, it will be useful if it were possible to estimate the reduction in variance for every test future test data. The challenge in such estimation is that the future data is not yet available and hence, unknown. Another related issue is the sensitivity to the numerical values

of the data points, i.e. by how much will the variance change if the same number of data points but a different set of values were used for calibration? For example, consider the simple case where Z is normally distributed with known variance σ^2. The mean needs to be estimated to be estimated based on realizations of Z. Conjugate distributions are available for this case; a normal prior is assumed for the mean and the posterior is also normal. For this special case, it is well-known [9] that the variance of the estimate of mean is independent of the numerical values of the realizations of Z, and is dependent only on the number of realizations of Z. It is easy to verify that this behavior cannot be generalized to all Bayesian updating scenarios. This chapter addresses the sensitivity of the model parameter to both the number and the numerical values of the data available for calibration.

The rest of this chapter is organized as follows. Section 9.2 discusses a preliminary approach for sensitivity analysis and identifies the difficulties in applying the methods for forward GSA to inverse sensitivity analysis. Section 9.3 revisits Bayesian calibration under uncertainty and proposes an efficient methodology to estimate the marginal PDFs using Bayesian updating. Section 9.4 develops the methodology for sensitivity analysis of parameter estimation. Section 9.5 extends the proposed methodology to multi-level systems, where multiple models and corresponding data are used to calibrate common model parameters; the sensitivity of parameters to (i) the uncertainties in all the models, and (ii) multiple levels of data are both quantified. The proposed methodology is illustrated using two numerical examples, first using a single-level model (Section 9.6), and then using multiple levels of models (Section 9.7).

9.2 Preliminary Approach

The effect of data on the variance of the parameter estimate can be studied in a relatively easy manner. This issue is directly discussed later in Section 9.4. On the

other hand, apportioning the variance of θ^i to other existing sources of uncertainty needs is more challenging. First, a preliminary approach is discussed in this section, and the forthcoming sections gradually develop a rigorous methodology. The other sources of uncertainty, as explained in Section 9.1, include both the uncertainty in the inputs, and the prior uncertainty in the other parameters (θ^{-i}). The Bayesian inference method, as shown in Eq. 9.1 calculates only the joint density of θ. This joint density is then used to calculate the marginal density and the variance of θ^i. A preliminary approach for inverse sensitivity analysis is to treat each of the other uncertain quantities to be deterministic (say, at the mean value), and estimate the new posterior variance of θ^i. Such a preliminary approach has been implemented for forward sensitivity analysis [179]; there is a general consensus that this preliminary approach is not accurate because (i) it is local since the uncertain quantity is fixed to be deterministic at a particular value; and (ii) the choice of the deterministic value significantly affects the posterior variance of θ^i. Hence, the GSA methodology was preferred for forward sensitivity analysis.

The next challenge is to apply the global sensitivity analysis methodology (discussed earlier in Section 2.6 in Chapter II) to inverse sensitivity analysis as well. In fact, it is quite easy to calculate the right-hand side expression in Eq. 2.15, i.e. the calculation of first-order index. Hence, it would seem as if the first-order index of an input X^j is simply equal to $\frac{V(E(\theta^i|X^j))}{V(\theta^i)}$. However, this is not true.

As explained earlier in Section 2.6, the fundamental theorem that governs the calculation of forward GSA indices is the variance decomposition theorem, i.e. $V(Y) = V(E(Y|X^j)) + E(V(Y|X^j))$, for all j. This was true in forward GSA because there was a deterministic transfer function ($Y = G(\boldsymbol{X})$) from the inputs $\boldsymbol{X}$ to the output Y. However, in the case of inverse sensitivity analysis, while estimating the variance of a particular model parameter, θ^i, $V(\theta^i) \neq V(E(\theta^i|X^j)) + E(V(\theta^i|X^j))$. This is

because (1) the parameter θ^i cannot be expressed as a deterministic function of all the inputs $\boldsymbol{X}$; (2) the variance of θ^i also depends on the prior PDFs of θ^{-i}; and (3) more interestingly, even when the inputs are deterministic, and there is only one parameter θ to be estimated, the Bayesian calibration procedure leads to a non-deterministic (uncertain) θ. Hence, there is no deterministic transfer function from $\boldsymbol{X}$ to $\boldsymbol{\theta}$ in order to facilitate variance decomposition; therefore, Eq. 2.15-2.17 are not applicable for inverse sensitivity analysis.

Sections 9.3 and 9.4 develop a new perspective on Bayesian calibration where a deterministic transfer function from $\boldsymbol{X}$ to $\boldsymbol{\theta}$ is developed. This deterministic transfer function is then used to guide in inverse sensitivity analysis.

9.3 Bayesian Calibration: New Perspective

Consider the computational model $Y = G(\boldsymbol{X}; \boldsymbol{\theta})$; though this model is deterministic (the output Y is single-valued for a given realization of inputs $\boldsymbol{X}$ and parameters $\boldsymbol{\theta}$), the inputs $\boldsymbol{X}$ are uncertain with probability density function (PDF) given by $f_{\boldsymbol{X}}(\boldsymbol{x})$, and the parameters $\boldsymbol{\theta}$ are unknown. Data D is collected on the output, and by assuming that data is unbiased ($D = Y + \epsilon$ where ϵ is $N(0, \sigma^2)$), the PDF of the parameters can be calculated using Bayes theorem as:

$$f''(\boldsymbol{\theta}) = \frac{f'(\boldsymbol{\theta})L(\boldsymbol{\theta})}{\int f'(\boldsymbol{\theta})L(\boldsymbol{\theta})d\boldsymbol{\theta}} \tag{9.1}$$

In Eq. 9.1, $f'(\boldsymbol{\theta})$ and $f''(\boldsymbol{\theta})$ refer to the prior and the posterior PDF of the parameters, and $L(\boldsymbol{\theta})$ is the likelihood function, which is proportional to the probability of the observing the given data D conditioned on $\boldsymbol{\theta}$. The denominator in Eq. 9.1 is simply a normalizing constant to ensure that the area under the posterior PDF is equal to unity.

The likelihood function needs to account for the uncertainty in the inputs, as:

$$L(\boldsymbol{\theta}) = \int f(D|y) f_Y(y|\boldsymbol{x};\boldsymbol{\theta}) f_{\boldsymbol{X}}(\boldsymbol{x}) d\boldsymbol{x} \tag{9.2}$$

Further, the likelihood function is always meaningful only upto a proportionality constant; sometimes, for the sake of simplicity, the proportionality is replaced with equality because when the likelihood function is substituted into Eq. 9.1, the normalizing constant will account for this proportionality.

The above likelihood function accurately accounts for all the uncertainty in the inputs. The calculation of this likelihood function requires forward uncertainty propagation (propagating the input uncertainty for a given realization of $\boldsymbol{\theta}$). If Monte Carlo sampling is used for this purpose, the number of function evaluations is equal to N_{MCS} (usually of the order of 10^3 or 10^4). Once the likelihood function is calculated, it is substituted into Eq. 9.1 for posterior evaluation. Markov Chain Monte Carlo (MCMC) sampling is then used to generate posterior samples of $\boldsymbol{\theta}$. If the corresponding number of function evaluations is equal to N_{MCMC} (usually of the order of 10^4 or 10^5), then the entire Bayesian updating procedure requires $N_{MCS} \times N_{MCMC}$. This approach is not only computationally expensive, but also combines all the uncertainty in the inputs in the likelihood function, and hence cannot provide a deterministic transfer function from $\boldsymbol{X}$ to $\boldsymbol{\theta}$.

9.3.1 The Conditional Posterior

This section proposes a new approach in which the marginal distribution for each parameter θ^i can not only be obtained in a computationally efficient manner but also facilitates a transfer function from $\boldsymbol{X}$ to $\boldsymbol{\theta}$ to aid in inverse sensitivity analysis. Earlier, it was stated that Eq. 9.2 calculates the likelihood function, by including the uncertainty in the inputs. Alternatively, the likelihood function can also be calculated

for a particular realization of the inputs; this likelihood is referred to as the conditional likelihood and is calculated as:

$$L(\boldsymbol{\theta};\boldsymbol{x}) = \int f(D|y) f_Y(y|\boldsymbol{x};\boldsymbol{\theta}) dy \tag{9.3}$$

Since $Y = G(\boldsymbol{X};\boldsymbol{\theta})$ is deterministic for given realization of $\boldsymbol{X}$ and $\boldsymbol{\theta}$, this conditional likelihood is simply equal to:

$$L(\boldsymbol{\theta};\boldsymbol{x}) = \frac{1}{\sqrt{2\pi}\sigma} exp(-\frac{(D-y)^2}{2\sigma^2}) \tag{9.4}$$

Once the conditional likelihood is calculated, it can be used in Bayes theorem to calculate the conditional posterior as:

$$f''(\boldsymbol{\theta};\boldsymbol{x}) = \frac{f'(\boldsymbol{\theta}) L(\boldsymbol{\theta};\boldsymbol{x})}{\int f'(\boldsymbol{\theta}) L(\boldsymbol{\theta};\boldsymbol{x}) d\boldsymbol{\theta}} \tag{9.5}$$

Thus, the conditional posterior is the posterior PDF calculated for a particular realization of the uncertain quantity $\boldsymbol{X}$. The true (unconditional) posterior, in turn, can be calculated from the conditional posterior as:

$$f''(\boldsymbol{\theta}) = \int f''(\boldsymbol{\theta};\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}) d\boldsymbol{x} \tag{9.6}$$

It can be seen that the difference between the conventional approach (Eq. 9.2 and Eq. 9.1) and conditional likelihood approach (Eqs. 9.5 – 9.6) is that the integration over the domain of $\boldsymbol{X}$ has simply been postponed to a later stage; both the approaches are mathematically equivalent.

9.3.2 Marginal Distribution and Variance

In order the perform inverse sensitivity analysis, it is essential to calculate the marginal distribution of every parameter (θ^i) from the joint posterior PDF $f''(\boldsymbol{\theta})$ by integrating over the domain of $\boldsymbol{\theta^{-i}}$. The marginal posterior of θ^i does not depend on the posterior of any other parameter ($\boldsymbol{\theta^{-i}}$). However, the prior of $\boldsymbol{\theta^{-i}}$, i.e. $f'(\boldsymbol{\theta^{-i}})$ directly affects the marginal posterior of θ^i. The marginal PDF of θ^i can be calculated as:

$$f''(\theta^i) = \frac{f'(\theta^i)L(\theta^i)}{\int f'(\theta^i)L(\theta^i)d\theta^i} \tag{9.7}$$

In Eq. 9.7, the true likelihood function (that includes the uncertainty in $\boldsymbol{x}$ and $\boldsymbol{\theta^{-i}}$) needs to be used; this is calculated as:

$$L(\theta^i) = \int f(D|y)f_Y(y|\boldsymbol{x};\boldsymbol{\theta^{-i}})f_{\boldsymbol{X}}(\boldsymbol{x})f'(\boldsymbol{\theta^{-i}})d\boldsymbol{x}d\boldsymbol{\theta^{-i}} \tag{9.8}$$

Thus, it is clear that, with respect to θ^i, $\boldsymbol{\theta^{-i}}$ can be treated as simply an additional source of uncertainty, similar to $\boldsymbol{X}$. Now, implementing the conditional likelihood method, the marginal, conditional posterior of θ^i can be calculated as:

$$f''(\theta^i;\boldsymbol{x},\boldsymbol{\theta^{-i}}) = \frac{f'(\theta^i)L(\theta^i;\boldsymbol{x},\boldsymbol{\theta^{-i}})}{\int f'(\theta^i)L(\theta^i;\boldsymbol{x},\boldsymbol{\theta^{-i}})d\theta^i} \tag{9.9}$$

The conditional, marginal likelihood (i.e. $L(\theta^i;\boldsymbol{x},\boldsymbol{\theta^{-i}})$) is calculated exactly as in the right hand side of Eq. 9.4.

9.3.3 Interpreting the Marginal Conditional Posterior

Consider the marginal conditional posterior (MCP) given by $f''(\theta^i;\boldsymbol{x},\boldsymbol{\theta^{-i}})$. Evidently, every choice of $\boldsymbol{x}$ and $\boldsymbol{\theta^{-i}}$ leads to a different MCP. In order obtain the

marginal unconditional posterior, an integration similar to Eq. 9.6 is performed, as:

$$f''(\theta^i) = \int f''(\theta^i; \boldsymbol{x}, \boldsymbol{\theta^{-i}}) f_{\boldsymbol{X}}(\boldsymbol{x}) f'(\boldsymbol{\theta^{-i}}) d\boldsymbol{x} d\boldsymbol{\theta^{-i}} \tag{9.10}$$

Note that this integration is similar to the calculation of the predictive posterior distribution (in the context of distribution parameter uncertainty), which was discussed in detail in Section 3.4. Therefore, Eq. 9.10 is analogous to Eq. 3.8 in Section 3.4. Therefore, $\boldsymbol{x}$ and $\boldsymbol{\theta^{-i}}$ can be viewed as distribution parameters of the conditional posterior $f''(\theta^i; \boldsymbol{x}, \boldsymbol{\theta^{-i}})$. Hence, to generate samples from the marginal unconditional posterior ($f''(\theta^i)$), a single loop sampling algorithm, similar to that in Section 3.4, is proposed:

1. Select one random sample of $\boldsymbol{X}$ and $\boldsymbol{\theta^{-i}}$ from $f_{\boldsymbol{X}}(\boldsymbol{x})$ and $f'(\boldsymbol{\theta^{-i}})$ respectively.

2. Use Bayesian updating to calculate the MCP $f''(\theta^i; \boldsymbol{x}, \boldsymbol{\theta^{-i}})$ for above $\boldsymbol{X}$ and $\boldsymbol{\theta^{-i}}$

3. Generate one random sample from the above MCP using CDF inversion (by generating a sample from the uniform distribution [0, 1]). This one sample represents a random realization from the marginal unconditional posterior $f''(\theta^i)$

4. Repeat above steps multiple times to obtain multiple samples of θ^i. These samples can be used to construct the marginal unconditional posterior $f''(\theta^i)$.

9.3.4 Computing the Marginal Conditional Posterior

Note that the computation of the marginal conditional posterior (MCP) for a given choice of $\boldsymbol{x}$ and $\boldsymbol{\theta^{-i}}$ requires only a single dimension integration as against the multi-dimensional integration needed for joint posterior in Eq. 9.1. Hence, the MCP can be evaluated by simple quadrature-based numerical integration techniques.

Such a technique was developed in Section 4.3.7 to replace MCMC sampling when the number of dimensions is small.[1] Since, the computation of MCP requires only one-dimensional integration, this techniques can be used readily.

The use of Laplace integration provides a computationally efficient alternative to numerical integration. Researchers [240, 241] have used this method to approximate the entire posterior distribution. However, this procedure unreasonably assumes that the joint posterior of all the calibration parameters is a multi-variate normal PDF without any dependence between the individual parameters. On the other hand, it may be reasonable to assume that the marginal conditional posterior is normally distributed, when the prior is uniform and the marginal conditional likelihood function is calculated using a Gaussian kernel.

In fact, this approach, i.e. MCP-based Laplace integration, may be applicable to all Bayesian updating problems when the focus is to estimate the marginal posterior distributions of the parameters. The correlations and dependencies are not ignored in this procedure; they are not calculated because this method directly calculates the marginal posterior rather than the joint posterior. Further research is necessary to explore and demonstrate the applicability of this method to Bayesian updating problems, in general. This will be considered in future work as this chapter focuses only on inverse sensitivity analysis.

9.4 Sensitivity Analysis Methodology

This section develops the computational methodology for inverse sensitivity analysis. The sensitivity of the estimate model parameter to both the available data, and the additional sources of uncertainty are considered.

[1]Such an integration technique has been used for the inference of equivalent initial flaw size in fatigue crack growth analysis [181, 199]

9.4.1 Sensitivity to Data

As stated earlier, the variance of a particular parameter θ^i depends on (1) number of data points; and (2) location of data points. Suppose that the available data points have been used to estimate θ^i. To quantify the effect of the number of data points, a three step procedure can be followed.

1. Based on the available data points, use bootstrapping to select random candidate for one additional data point.

2. Estimate updated variance of θ^i by considering both old points and the new candidate data point, and hence, estimate the change in variance.

3. Repeat the above steps by considering multiple random candidates and calculate the expectation of change in variance.

To quantify the effect of location of data points, a similar three step procedure can be followed:

1. Based on the available data points, use bootstrapping to select random candidates for data points equal in number to the available data points.

2. Replace original data points with the new set and estimate θ. Calculate the difference in variance.

3. Repeat the above steps by considering multiple random sets and calculate the expectation of change in variance.

9.4.2 Sensitivity to Sources of Uncertainty

To apply variance-based global sensitivity indices, it was explained earlier in Section 9.2 that a deterministic transfer function from $\boldsymbol{X}$ to $\boldsymbol{\theta}$. Further, it was also

discussed that a particular model parameter θ^i is sensitive to both $\boldsymbol{X}$ and $\boldsymbol{\theta^{-i}}$. Hence, the deterministic transfer function should be of the form $\theta^i = H(\boldsymbol{X}, \boldsymbol{\theta^{-i}})$. The marginal conditional posterior PDF discussed in Section 9.3 is used to construct this deterministic transfer function H.

For a given choice of $\boldsymbol{x}$ and $\boldsymbol{\theta^{-i}}$, the estimate of θ is <u>not</u> deterministic and the uncertainty is given by the MCP $f''(\theta^i; \boldsymbol{x}, \boldsymbol{\theta^{-i}})$. This uncertainty, according to the Bayesian philosophy, is a subjective estimate of uncertainty in θ^i for the chosen $\boldsymbol{x}$ and $\boldsymbol{\theta^{-i}}$. The lack of a deterministic function prohibits the application of the existing global sensitivity analysis methodology. A new auxiliary variable U_{θ^i} is introduced to represent the aforementioned Bayesian subjective uncertainty; this uncertainty is a result of the "subjectivity" inherent in the Bayesian philosophy for inference. This auxiliary variable is defined such that:

$$\theta^i = \int_{-\infty}^{U_{\theta^i}} f''(\theta^i; \boldsymbol{x}, \boldsymbol{\theta^{-i}}) d\theta^i \tag{9.11}$$

Note that U_{θ^i} is simply equal to the CDF value of the marginal conditional posterior $f''(\theta^i; \boldsymbol{x}, \boldsymbol{\theta^{-i}})$, and hence uniformly distributed on the interval [0, 1]. Refer to the algorithm to generate samples from the marginal unconditional posterior, discussed earlier in Section 9.3.3. In step 3, a random number (uniformly distributed on [0, 1]) was chosen to draw one sample from the MCP. The auxiliary variable is exactly same as this random number.

Using this auxiliary variable, Eq. 9.11 can be rewritten to define a deterministic function, as follows:

$$\theta^i = H(\boldsymbol{x}, \boldsymbol{\theta^{-i}}, U_{\theta^i}) \tag{9.12}$$

The sampling algorithm (to generate samples of θ^i from its marginal unconditional PDF) described in Section 9.3.3 is equivalent to performing Monte Carlo sampling on

Eq. 9.12 by generating samples of $\boldsymbol{x}$, $\boldsymbol{\theta^{-i}}$, and U_{θ^i} simultaneously. Further, Eq. 9.12 provides a deterministic function to facilitate inverse global sensitivity analysis. The uncertainty in each model parameter θ^i can be apportioned to the other inputs, other model parameters, and the subjectivity term.

The first-order effects index with respect to the contribution of a particular input (x^j) to the model parameter (θ^i) can be calculated similar to Eq. 2.15 (by fixing x^j alone) as:

$$S_{1,i}^{x^j} = \frac{V(E(\theta^i|x^j))}{V(\theta^i)} \tag{9.13}$$

Similarly, the total effects index with respect to the contribution of a particular input (x^j) to the model parameter (θ^i) can be calculated similar to Eq. 2.17 (by fixing all quantities other than x^j) as:

$$S_{T,i}^{x^j} = 1 - \frac{V(E(\theta^i|x^{-j}, \boldsymbol{\theta^{-i}}, U_{\theta^i}))}{V(\theta^i)} \tag{9.14}$$

Similarly, the first-order effects and total effects indices with respect to the contribution of a particular model parameter (θ^j) to the model parameter (θ^i where $i \neq j$) can be calculated as:

$$S_{1,i}^{\theta^j} = \frac{V(E(\theta^i|\theta^j))}{V(\theta^i)} (i \neq j) \tag{9.15}$$

$$S_{T,i}^{\theta^j} = 1 - \frac{V(E(\theta^i|\boldsymbol{x}, \theta^{-(i,j)}, U_{\theta^i}))}{V(\theta^i)} (i \neq j) \tag{9.16}$$

Further, the contribution of the subjectivity term (U_{θ^i}) can be calculated as:

$$S_{1,i}^{U_{\theta^i}} = \frac{V(E(\theta^i|U_{\theta^i}))}{V(\theta^i)} \tag{9.17}$$

$$S_{T,i}^{U_{\theta^i}} = 1 - \frac{V(E(\theta^i|\boldsymbol{x}, \boldsymbol{\theta^{-i}}))}{V(\theta^i)} \tag{9.18}$$

Thus, the concept of marginal conditional posterior has been used to derive a methodology for inverse sensitivity analysis analogous to the existing approach for forward sensitivity analysis. In the next section, the concept of inverse sensitivity analysis is extended to multi-level systems, where the same set of model parameters are calibrated using multiple models and corresponding data sets.

9.5 Extension to Multi-level Calibration

Consider a model parameter which needs to be calibrated using multiple models and data sets; without loss of generality, the method is discussed for two different models (G_1 and G_2) and corresponding data (D_1 and D_2). These two models have their own sources of uncertainty ($\boldsymbol{x_1}$ and $\boldsymbol{x_2}$) described in terms of the PDFs ($f_{X_1}(\boldsymbol{x_1})$ and $f_{X_2}(\boldsymbol{x_2})$ respectively).

$$Y_1 = G_1(\boldsymbol{\theta}, \boldsymbol{x_1}) \tag{9.19}$$

$$Y_2 = G_2(\boldsymbol{\theta}, \boldsymbol{x_2}) \tag{9.20}$$

As discussed in earlier chapter, the concept of Bayesian network [25, 26] is used to connect multiple models, corresponding data sets, and the various sources of uncertainty. As explained earlier in Section 2.4.4, this Bayesian network is useful for uncertainty propagation as well as calibration of model parameters using multiple sets of data. If a prior PDF ($f'(\boldsymbol{\theta})$) is assumed for the model parameters ($\boldsymbol{\theta}$), then the likelihood function ($L(\boldsymbol{\theta})$) is calculated as being proportional to the probability of observing all data (both D_1 and D_2 in this case) conditioned on the model parameters ($\boldsymbol{\theta}$). Then, the posterior PDF of the model parameters is calculated using Eq. 9.1. The likelihood function is calculated as:

$$L(\boldsymbol{\theta}) \propto P(D_1, D_2|\boldsymbol{\theta}) \tag{9.21}$$

The traditional method for likelihood calculation includes the uncertainty in x_1 and x_2, as:

$$L(\boldsymbol{\theta}) = \int L(\boldsymbol{\theta}|x_1, x_2) f_{X_2}(\boldsymbol{x_2}) f_{X_1}(\boldsymbol{x_1}) dx_1 dx_2 \tag{9.22}$$

In Eq. 9.22, $L(\boldsymbol{\theta}|x_1, x_2)$ refers to the conditional likelihood (which is proportional to $P(D_1, D_2|\boldsymbol{\theta}, x_1, x_2)$). Assuming that data is unbiased ($D_1 = y_1 + \epsilon_1$ where ϵ_1 is $N(0, \sigma_1^2)$ and $D_2 = y_2 + \epsilon_2$ where ϵ_2 is $N(0, \sigma_2^2)$), and the data are collected independently, the conditional likelihood is calculated as:

$$L(\boldsymbol{\theta}|x_1, x_2) \propto \frac{1}{\sqrt{2\pi}\sigma_1} exp(-\frac{(D_1 - y_1)^2}{2\sigma_1^2}) \frac{1}{\sqrt{2\pi}\sigma_2} exp(-\frac{(D_2 - y_2)^2}{2\sigma_2^2}) \tag{9.23}$$

Then the likelihood function $L(\boldsymbol{\theta})$ is substituted into Eq. 9.1 to calculate the joint posterior of the model parameters ($\boldsymbol{\theta}$).

For the purpose of inverse sensitivity analysis, the concept of marginal conditional posterior introduced in Section 9.3 can be extended to the case of multi-level calibration as well, and the MCP $f''(\theta^i|\boldsymbol{x_1}, \boldsymbol{x_2}, \boldsymbol{\theta^{-i}})$ are estimated. The sampling algorithm developed in Section 9.3.3 is then used to calculate the marginal unconditional posterior $f''(\theta^i)$ and hence, the posterior variance.

The posterior variance is used to calculate the sensitivity to the calibration data (both number and location). This analysis is useful in resource allocation; one of D_1 or D_2 may be more influential in reducing the variance of a particular parameter θ^i, and the results of sensitivity analysis aids in such decision-making.

The MCP $f''(\theta^i|\boldsymbol{x_1}, \boldsymbol{x_2}, \boldsymbol{\theta^{-i}})$, along with the auxiliary variable U_{θ^i} can be used to analyze the sensitivity of other uncertain inputs ($\boldsymbol{x_1}$ and $\boldsymbol{x_2}$), other model parameters ($\boldsymbol{\theta^{-i}}$) to θ^i. The following sections present numerical examples to illustrate the proposed inverse sensitivity analysis methodology; Section 9.6 discusses single-level calibration and Section 9.7 discusses multi-level calibration.

9.6 Numerical Example: Single-level Calibration

This numerical example deals with the estimation of equivalent initial flaw size (EIFS) in fatigue crack growth analysis in mechanical components. A rigorous approach to fatigue life would account for crack growth starting from the actual initial flaw, accounting for imperfections, voids and non-metallic inclusions. This procedure is cumbersome because small crack growth propagation is anomalous in nature and hence not completely understood. On the other hand, there are several crack growth models (Paris law [242], FASTRAN [243], AFGROW [244], etc.) in the long crack regime which are used to study long crack growth behavior. Equivalent initial flaw size is a fictitious quantity which was introduced to bypass small crack growth calculations and make direct use of a long crack growth law in order to make fatigue life prediction; the EIFS must be chosen in such a way that when the long crack growth law is used with EIFS as the initial value, it yields crack growth results that match with observed crack growth data [188].

Since EIFS is fictitious and hence, not measurable, it needs to be estimated based on observed data on crack size. Initially, back extrapolation techniques [245] were used for EIFS calculation, and then likelihood-based [180] and Bayesian [182] methods have been implemented for the statistical inference of EIFS. Chapter V presented a Bayesian methodology for estimating EIFS by considering complicated geometry, multi-axial variable amplitude loading, and multiple sources of uncertainty including physical variability, data uncertainty, and model errors.[1] Since the focus of the present chapter is on inverse sensitivity analysis, the present chapter considers a simpler problem, i.e. plate subjected to cyclic, uniform uniaxial stress (S), and Paris' law is used for crack growth analysis. Paris law calculates the increment in crack size

[1] Chapter V estimated all crack growth parameters, in addition to EIFS. The sole estimation of EIFS under complicated geometry and multi-axial variable amplitude loading has also been studied [181, 197, 199]

per cycle of loading, in terms of crack growth parameters (C and n), threshold stress intensity factor (ΔK_{th}), and load stress intensity factor (ΔK):

$$\frac{da}{dN} = C(\Delta K)^n(1 - \frac{\Delta K}{\Delta K_{th}})^p \tag{9.24}$$

The stress intensity factor (ΔK), for the sake of illustration, is assumed to be available in closed form, as:

$$\Delta K = S\sqrt{\pi a} \tag{9.25}$$

Data is collected in the form of "crack size (A) measured after a number of cycles (N)", and this data is used to estimate EIFS; 10 such data points are assumed to be available for calibration of EIFS. The prior of EIFS is chosen to be uniform over the region where $L(\theta) \neq 0$. Any other prior uniformly distributed over a wider range will lead to the same posterior PDF.

For the sake of illustration, the quantities $S \sim N(400, 40)$, $C \sim N(6.54 \times 10^{-13}, 6.54 \times 10^{-14})$, and $\Delta K_{th} \sim N(5.66, 0.6)$ are chosen to be random variables. The uncertainty in the equivalent initial flaw size significantly affects the uncertainty in the fatigue life prediction. Hence, it is important to analyze the sensitivity of equivalent initial flaw size, and thereby help in identifying ways to reduce this uncertainty. The prior PDF, the posterior PDF, and the marginal conditional PDF (i.e. MCP conditioned at the mean values of S, C, and ΔK_{th}) are shown in Fig. 9.1. The posterior variance is equal to 1.3×10^{-9}.

The sensitivities of both (1) location and number of data; and (2) additional sources of uncertainty to the equivalent initial flaw size are quantified. The sensitivity of location of data is estimated by recalculating the variance by considering multiple alternate sets of 10 data points. Hence, this variance is a random variable whose

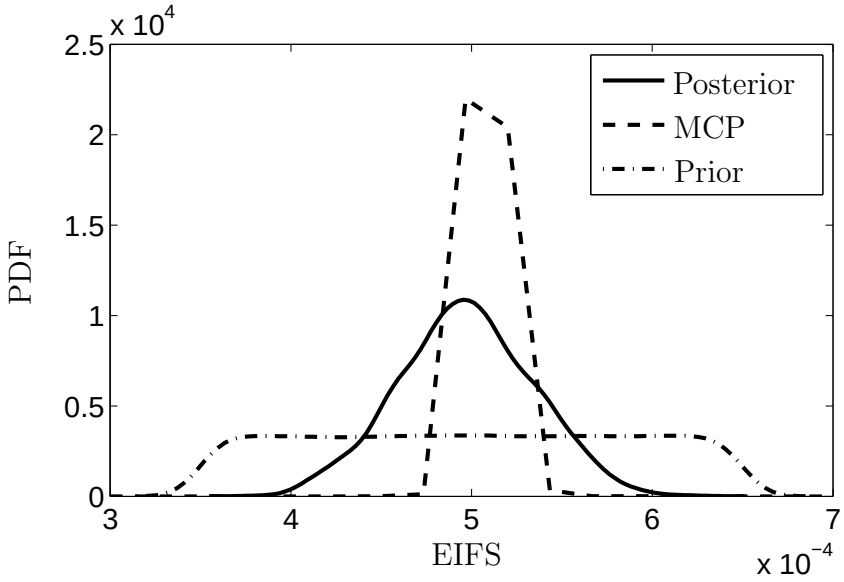


Figure 9.1: Calibration of EIFS

statistics are estimated to be mean (of variance) $= 1.3 \times 10^{-9}$ and standard deviation (of variance) $= 0.04 \times 10^{-9}$.

The sensitivity of number of data is estimate by generating several additional data points and calculating the expected value of variance, and percentage of the original posterior variance. The results are shown in Table 9.1. It can be seen from Table 9.1

Table 9.1: Sensitivity of EIFS to Number of Data Points

Number of Data Points	E(Variance of EIFS)	Percentage of Posterior Variance
10	1.30×10^{-9}	100.0%
11	1.25×10^{-9}	96.2%
12	1.20×10^{-9}	92.3%
13	1.10×10^{-9}	84.6%
14	1.07×10^{-9}	82.3%

that the expected value of variance of EIFS gradually decays with the number of data. The decrease is gradual and it is observed that the decrease is gradual after 14 tests.

The sensitivity of the different sources of uncertainty is calculated using the

marginal conditional posterior-based methodology and the first-order and total effects indices are shown in Table 9.2.

Table 9.2: Global Sensitivity of EIFS: Additional Uncertainty Sources

Source of Uncertainty	First-order Index	Total Effects Index
C	0.0	0.42
ΔK_{th}	0.10	0.12
S	0.59	0.65
Subjectivity (U_θ)	0.16	0.19

From Table 9.2, it is seen that the contribution of load uncertainty (S) to the uncertainty in EIFS is maximum. Though the first-order contribution of C is approximately equal to zero, it has a significant total effects contribution, indicating higher order interactions; higher order interactions are significant because Eq. 9.24 consists of the product of powers of the terms C, S, and ΔK_{th}.

The above sensitivity analysis provides a methodology using which it is possible to answer the question - "what causes uncertainty in the EIFS estimate?" If the goal is to reduce the uncertainty in EIFS, this can be established by either collecting data or by reducing the uncertainty in S, ΔK_{th}, or C, by the amounts indicated in Table 9.1 (through data collection) and Table 9.2 (reducing uncertainty).

9.7 Numerical Example: Multi-level Calibration

This section considers the multi-physics thermal vibration problem which was discussed earlier in Section 8.3. This problem is representative of the thermal vibration in solar arrays of space telescopes and spacecraft booms [208]. In Chapter VIII, the focus was to select tests in order to reduce the variance of the multi-disciplinary system reponse, i.e. the deflection of the vertical beam. In this chapter, the focus is on parameter estimation and inverse sensitivity analysis, when the data is collected

through two types of tests – flexural subsystem test and thermal subsystem test. Therefore, the equations that predict the flexural response (without thermal effects) and the thermal response (without flexural effects) are of interest in this chapter, in contrast with the multi-disciplinary response in Chapter VIII. The equations for the individual disciplines are explained in the forthcoming subsections. The numerical details regarding the inputs, parameters, material and geometric properties, including the calibration quantities, are explained later in Section 9.7.3.

9.7.1 Flexural Subsystem Test

In the flexural subsystem test, the vertical cantilever beam is subjected to free vibration, the amplitude of the tip mass is measured at a particular instant, and this information is used to calibrate the model parameters (damping and radius). Refer to Section 9.7.3 for numerical details, including calibration parameters.

The equation governing the displacement of tip mass (V) is calculated in terms damping coefficient (ξ), angular frequency (ω_0), gravitational constant (g), and length of beam (l), as:

$$\ddot{V} + 2\xi\omega_0\dot{V} + (\omega_0^2 + \frac{3g}{2l})V = 0 \tag{9.26}$$

The angular frequency (ω_0) depends on the elastic modulus (E), area moment of inertia (I), length of tube (l), tip mass (m), as:

$$\omega_0 = \frac{3EI}{ml^3} \tag{9.27}$$

where the moment of inertia (I) is calculated in terms of tube radius (r) and tube wall thickness (w) as:

$$I = \frac{\pi}{64}(2r^4 - (2r - 2w)^4) \tag{9.28}$$

9.7.2 Thermal Subsystem Test

In the thermal subsystem test, a thin walled tube is subjected to heat flux (S), and the temperature is measured; this information is used to calibrate the model parameters (absorptivity and radius). Refer to Section 9.7.3 for numerical details, including calibration parameters.

The mathematical model to calculate this temperature (T_m) as a function of time (t) is constructed in terms of absorptivity (α_s), incident heat flux (S), mass density (ρ), specific heat capacity (c), tube wall thickness (w), and an exponential rate parameter (τ), as:

$$T_m = \frac{1}{2}\frac{\alpha_s S \tau}{\rho c w}(1 - exp(-\frac{t}{\tau})) \tag{9.29}$$

where the exponential rate parameter (τ) is calculated in terms of thermal conductivity (k), tube radius (r), specific heat capacity (c), Stefan-Boltzmann constant (σ), emissivity (ϵ), mass density (ρ), tube wall thickness (w), absorptivity (α_s), and incident heat flux(S) as:

$$\frac{1}{\tau} = \frac{k}{\rho c r^2} + \frac{4\sigma\epsilon}{\rho c w}(\frac{1}{\pi}\frac{\alpha_s S}{\sigma\epsilon})^{0.75} \tag{9.30}$$

9.7.3 Numerical Details

The independent quantities described in the above subsections are tabulated in Table 9.3. The dependent quantities (ω_0, V, I, T_m, τ) can be calculated based on the above equations.

As seen in Table 9.3, there are three calibration parameters; the damping coefficient (ξ) is calibrated only using flexural subsystem test data, the absorptivity (α_s) is calibrated only using thermal subsystem test data, whereas the radius (r) is calibrated using test data on both the subsystems. Note that radius (r) is a geometric property and often may not be desired to be calibrated since it can be directly measured; in

Table 9.3: Numerical Details

Name	Symbol	Description	Numerical Value	Unit
Damping Coefficient	ξ	Parameter	Prior $\sim U(0, 0.02)$	No unit
Tube radius	r	Parameter	Prior $\sim N(1.7, 0.05)$	$\times 10^{-3} m$
Absorptivity	α_s	Parameter	Prior $\sim U(0.6, 1)$	No unit
Specific heat capacity	c	Uncertainty	$N(500, 50)$	$Jkg^{-1}K^{-1}$
Elastic modulus	E	Uncertainty	$N(193, 19.3)$	$\times 10^9 N/m^2$
Thermal conductivity	k	Uncertainty	$N(16, 1.6)$	$Wm^{-1}K^{-1}$
Emissivity	ϵ	Uncertainty	$U(0.75, 1)$	No unit
Heat flux	S	Uncertainty	$N(1000, 50)$	W/m^2
Tip mass	m	Constant	150	$\times 10^{-3} kg$
Tube wall thickness	w	Constant	80	$\times 10^{-6} m$
Mass density	ρ	Constant	7930	kg/m^3
Stefan-Boltzmann constant	σ	Constant	5.67×10^{-8}	$Js^{-1}m^{-2}K^{-4}$
Gravitational Constant	g	Constant	9.81	m/s^2
Length of beam	l	Constant	0.69	m

this numerical example, it is chosen as the calibration parameter because it affects the output of both the subsystems.

Five calibration data for each subsystem are assumed to be available, and this data is used to calibrate all the three parameters. The prior PDF, posterior PDF, and marginal conditional PDF (i.e. MCP, conditioned on the mean of all other uncertain quantities) are shown in the following figures. The corresponding posterior variances are 6.5×10^{-7}, 0.006, and 3.6×10^{-9} respectively.

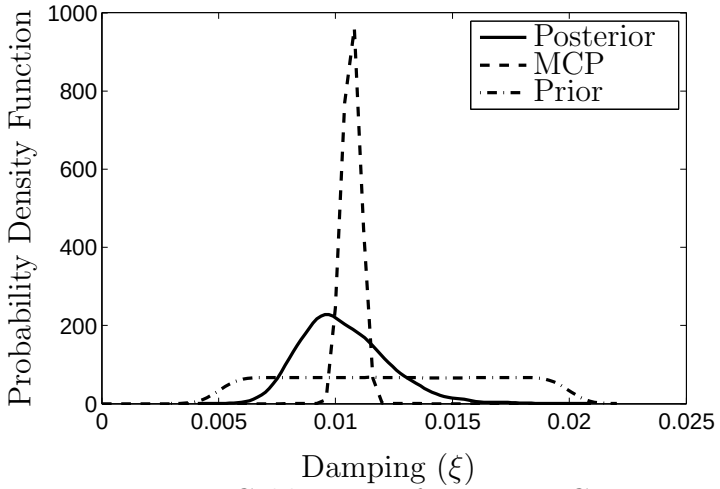


Figure 9.2: Calibration of Damping Constant

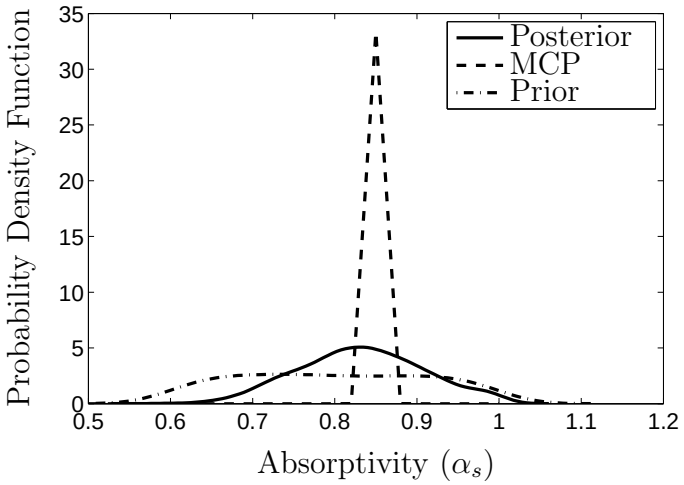


Figure 9.3: Calibration of Absorptivity

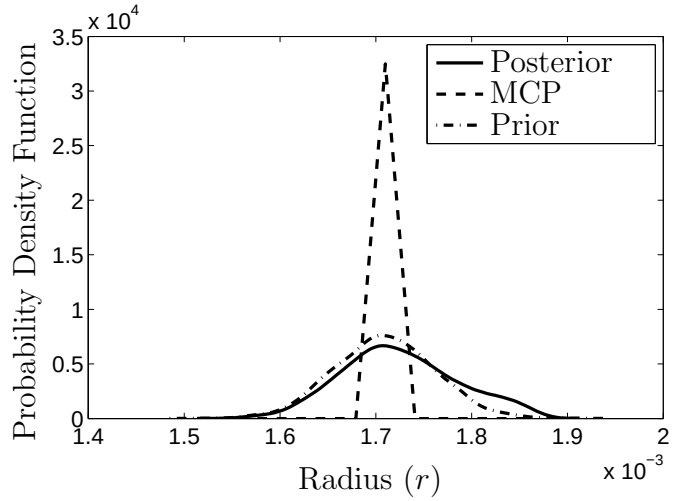


Figure 9.4: Calibration of Radius

The following sections present the results of inverse sensitivity analysis with respect to each of the three calibration parameters.

9.7.4 Sensitivity to Calibration Data

The sensitivity of the calibration parameter to both the location and number of test data points (flexural and thermal subsystem tests) are quantified, and tabulated in Table 9.4. Note that thermal subsystem test data cannot be used to infer the

damping (ξ) while the flexural subsystem test data cannot be used to infer absorptivity (α_s). In the case of inferring r, while computing the sensitivity of the data on the particular subsystem, the data on the other subsystem remains unchanged from the original calibration data (i.e. 5 experiments on each of the two subsystems).

Table 9.4: Sensitivity to Data

Calibration Parameter	Damping (ξ)	Absorptivity (α_s)	Radius (r)
Flexural Subsystem Data			
Sensitivity to Location			
Mean of Variance	6.5×10^{-7}	0.006	3.6×10^{-9}
Std of Variance	0.05×10^{-7}	0	0.8×10^{-9}
Sensitivity to Number : Percentage of Posterior Variance			
6^{th}	98 %	100 %	99.9 %
7^{th}	95 %	100 %	99.9 %
Thermal Subsystem Data			
Sensitivity to Location			
Mean of Variance	6.5×10^{-7}	0.006	3.6×10^{-9}
Std of Variance	0	0.0005	1.0×10^{-9}
Sensitivity to Number : Percentage of Posterior Variance			
6^{th} Data Point	100 %	95 %	100 %
7^{th} Data Point	100 %	93 %	99.9 %

9.7.5 Global Sensitivity to Sources of Uncertainty

The proposed inverse global sensitivity analysis methodology is used to apportion the uncertainty in the estimated model parameter to the other sources of uncertainty. Note that, for a particular calibration parameter, another calibration parameter is treated as an additional source of uncertainty described using its prior. The results of inverse sensitivity analysis are given in Tables 9.5 – 9.7.

Tables 9.5 – 9.7 indicate which quantities are the highest contributors of uncertainty to each of the calibration parameters. This analysis is useful when the aim is to reduce the uncertainty in the parameter estimate. The complete significance of

Table 9.5: Global Sensitivity of Damping: Additional Uncertainty Sources

Source of Uncertainty	First-order Index	Total Effects Index
c	0.0	0.0
E	0.40	0.63
k	0.0	0.0
α_s	0.0	0.0
S	0.0	0.0
r	0.32	0.55
ϵ	0.0	0.0
Subjectivity (U_ξ)	0.10	0.62

Table 9.6: Global Sensitivity of Absorptivity: Additional Uncertainty Sources

Source of Uncertainty	First-order Index	Total Effects Index
c	0.20	0.25
E	0.0	0.0
k	0.30	0.35
ϵ	0.0	0.01
S	0.28	0.28
r	0.10	0.11
ξ	0.0	0.0
Subjectivity (U_{α_s})	0.01	0.05

such estimation can be realized when computing the system-level deflection using the equations given by Thornton [208]. Higher uncertainty in the parameters would lead to higher uncertainty in the prediction. Thus, if the overall system-level prediction uncertainty needs to be decreased, the uncertainty in the parameters which are calibrated using lower-level tests also need to be decreased. The proposed methodology provides a quantitative approach for such problems, by studying the sensitivity of the calibration parameter to all the sources of uncertainty.

9.8 Conclusion

This chapter proposed a statistical methodology to study the sensitivity of model parameter estimation with respect to the various sources of uncertainty under which

Table 9.7: Global Sensitivity of Radius: Additional Uncertainty Sources

Source of Uncertainty	First-order Index	Total Effects Index
c	0.12	0.20
E	0.26	0.35
k	0.13	0.18
α_s	0.02	0.04
S	0.23	0.40
ξ	0.10	0.12
ϵ	0.05	0.08
Subjectivity (U_r)	0.05	0.10

the parameters need to be calibrated, and the calibration data using which the model parameters are estimated. The sensitivity to the data was computed with respect to both number of data points and the location of data points. In order to accomplish this, several realizations of data sets were simulated, and the posterior variance was calculated. The sensitivity to the various sources of uncertainty was estimated using the method of global sensitivity analysis. Until now, the method of global sensitivity analysis has been applied only to forward problems involving uncertainty propagation. This chapter extended this methodology to the inverse problem of model parameter estimation. The concept of marginal conditional posterior was introduced, and the effect of subjectivity inherent in the Bayesian philosophy was also quantified.

Future work needs to directly use the results of inverse sensitivity analysis in test design; for example, if the calibration is more sensitive to a particular uncertain quantity, then it may be necessary to consider several realizations of that particular quantity while designing tests and vice-versa. Other possible directions for further research are as follows: (1) The marginal conditional posterior approach seems to provide a computationally efficient alternative to compute the marginal posteriors of model parameters; this procedure needs to be further explored to check its applicability to generic Bayesian updating problems. (2) Further, this chapter considered only physical variability in inverse sensitivity analysis; future work needs to consider data

uncertainty, and model uncertainty. (3) If calibration needs to be performed under both aleatory and epistemic uncertainty, then the contribution of each of those to the overall estimation uncertainty can be calculated using the proposed approach. This analysis helps to quantify what extent of estimation uncertainty is reducible vs. what is irreducible.

CHAPTER X

CONCLUSION

10.1 Summary of Accomplishments

A critical requirement for the analysis and design of engineering systems is the ability to predict the performance of the system with a certain level of confidence. This is not trivial because the system performance is affected by several sources of uncertainty (physical variability, data uncertainty, and model uncertainty). The overall objective of this dissertation is to develop a unified framework for representation of uncertainty quantification, and seamlessly integrate the various sources of uncertainty across multiple models and compute the uncertainty in the system-level response, in order to provide information for decision-making in engineering systems.

The various accomplishments of this dissertation can be classified into two broad categories: (1) uncertainty quantification; and (2) uncertainty integration. In uncertainty quantification, the objective is to develop a framework for quantifying data uncertainty and model uncertainty. In uncertainty integration, multiple models (component-level, subsystem-level, and system-level models) are considered, and the goal is to compute the uncertainty in system-level prediction by integrating the various sources of uncertainty across multiple models and the different types of experimental data available across multiple levels.

First, this dissertation developed a likelihood-based approach for the probabilistic representation of data uncertainty due to sparse and interval data. This method was first developed for a random variable with known distribution type, and the uncertainty in the distribution parameters was quantified. Then, the likelihood-based

approach was extended to random variables with unknown distribution type; a parametric approach (by considering multiple competing distribution types) as well as a non-parametric approach were developed. For the parametric approach, a global sensitivity analysis-based methodology was developed to quantify the individual contributions of distribution type uncertainty, distribution parameter uncertainty and variability.

The second major accomplishment of this dissertation was the development of methods for the quantification of model uncertainty. Two types of solution approximation errors (discretization error and surrogate model error) were computed during model verification. The Bayesian approach for model calibration and two computational methods (Bayesian hypothesis testing and model reliability approach) for model validation were extended to account for the different sources of uncertainty, and include different types of data scenarios such as interval data, partially characterized data, and time series data.

The methods for quantifying data uncertainty and model uncertainty were applied to the problem of fatigue crack growth analysis, as a case study. While several sources of data uncertainty and model uncertainty had been ignored in the literature pertaining to fatigue crack growth, this dissertation accounted for all the sources of uncertainty for fatigue crack growth prediction and confidence assessment.

The third major accomplishment of this dissertation was with respect to uncertainty integration. A Bayesian network-based methodology was developed to integrate the results from various uncertainty quantification activities such as model verification, validation, and calibration in order to quantify the uncertainty in the overall system-level prediction. This methodology was developed for two different types of hierarchical systems (sequential and non-sequential).

The third type configuration includes feedback coupling amongst models. A new

methodology was developed for uncertainty propagation in multi-disciplinary systems with feedback coupling, and feedback coupling was replaced with equivalent unidirectional coupling, thereby transforming this hierarchy into a sequential system, and thus facilitating the use of the Bayesian network.

While the previous accomplishments with respect to uncertainty integration focused on quantifying the uncertainty in the system-level prediction, the fifth accomplishment was the development of a new methodology for test resource allocation, in order to achieve a reasonable reduction in the variance of the system performance prediction. This was achieved through calibration and uncertainty reduction in the parameters, thereby leading to uncertainty reduction in the system-level prediction.

The presence of the different sources and types of uncertainty have a profound influence on parameter uncertainty and system-level uncertainty reduction. Hence, in order to reduce the uncertainty in the parameters, it would be beneficial to know the sensitivity of these parameters to the various sources of uncertainty and the calibration data itself. Sometimes, a parameter may be more sensitive to one type of test vs. another type of test. The final accomplishment was the development of a new "inverse sensitivity methodology" in order to quantify the sensitivity of the model parameters. This was achieved through the use of the concepts of "marginal conditional posterior" and global sensitivity analysis.

10.2 Future Work : Short Term

The topics of uncertainty quantification and integration are fertile and there are several possible directions for further research.

This dissertation only addressed data uncertainty due to sparse point and interval

data. Sometimes, information may be available in the form of qualitative or categorical information. Future work needs to address these situations and develop methods for propagating such information in a probabilistic framework.

The Bayesian hypothesis testing and model reliability metric approaches for model validation are challenging from the perspective of numerical implementation. The former requires the choice of an alternate hypothesis which significantly affects the validation metric, and hence is not an absolute measure. While the model reliability metric is absolute, it may tend to zero as the number of validation data increases. Future work needs to overcome this challenge and propose validation metrics that are not only meaningful, but also facilitate the calculation of the probability that the model is correct, since this probability is instrumental in the integration of results from model verification, validation, and calibration. Further, this methodology of integration needs to be extended to multi-disciplinary and multi-scale systems.

With respect to uncertainty quantification in multi-disciplinary analysis, there are two directions for future work. (1) The proposed likelihood-based method was demonstrated only for a small number of coupling variables. Future work needs to address situations when the coupling is a field quantity, for e.g., pressures and displacements in fluid-structure interaction. (2) One advantage of the proposed method was that the feedback coupling between models can be replaced with uni-directional coupling. Therefore, the Bayesian network, which is acyclic in nature, can now be used to facilitate calibration, validation and integration in multi-disciplinary systems. Though this capability was mentioned in Chapter VII, the methodology needs to be developed and applied to practical problems in future work.

The optimization methodology for test resource allocation answered two questions: (1) What type of test to perform? (2) How many tests of each type? Future work needs to address design of tests by answering the question: At what settings should the

tests be performed? The proposed inverse sensitivity analysis methodology should be used to guide the test design. Also, this methodology needs to be extended to include other sources of uncertainty, and quantify the contributions of aleatory and epistemic uncertainty to model parameter estimation.

10.3 Future Work: Long Term

While most of the aforementioned objectives are short term goals, it may be beneficial to identify a set of long-term research goals and needs in order to advance the state of the art in the area of uncertainty quantification.

10.3.1 Uncertainty Quantification

Though the basic methods of uncertainty quantification such as Monte Carlo simulation, Bayesian updating, reliability methods, etc. are well established in the literature, their application to practical problems is quite daunting. Contemporary systems are often multi-disciplinary with multiple levels of modeling, and multiple scales of analysis. Future work needs to develop computational methods that can address uncertainty quantification across multiple scales and multiple physics of modeling; this is challenging because it is necessary to rigorously account for the different sources of uncertainty while propagating information from one scale and/or discipline to another.

10.3.2 Information Fusion

Information is available in various forms, such as models, test data, expert opinion, etc. A challenging task is to integrate all the available sources of information

for the purpose of uncertainty quantification. One hurdle is that different types of information are currently treated using different mathematical methods. Future work needs to develop a a unified, integrated methodology by assimilating information from multiple sources and help in meaningful, risk-informed decision-making.

10.3.3 Decision-making under Uncertainty

There are various aspects of decision-making during different stages of the life-cycle of an engineering system. Future work needs to develop methods for decision-making in all of these stages by rigorously accounting for the various types of uncertainty. In the analysis stage, model selection and development, verification, validation, and calibration are of interest. Currently, several of these activities are performed individually, and there is an evident need for integration of these activities during the analysis stage. Such integration can alone provide a rigorous measure of uncertainty, and aid in the overall risk-assessment. The design stage comprises of both system design and test design. Fundamentally, both these problems are optimization problems. While the former has been discussed more extensively in the literature, the latter has not yet received significant attention. In the operations stage, the focus is on structural health monitoring, which by itself is a large topic.

10.3.4 Structural Health Monitoring

Once the engineering system is in use, it is necessary to not only constantly monitor its health, but also provide prognosis, thereby to calculate the remaining useful life of the system. The field of structural health monitoring has gained considerable attention over the past several years; however, in recent times, the importance of

quantifying the uncertainty in diagnosis[1] and prognosis[2]. Future work needs to develop methods to quantify the confidence in diagnosis and prognosis, by including multiple levels of data and models, and by rigorously accounting for the different types of uncertainty.

10.3.5 Computational Efficiency

All of the above efforts require large computational resources since several hundreds of thousands of computational model evaluations may be necessary. One simple solution is to replace computational models with surrogate models (polynomial chaos expansions, Gaussian processes, support vector and relevance vector machines, etc.). Another solution is to focus on the mathematical foundations of uncertainty quantification and find intelligent ways to achieve a substantial decrease in computational effort. The importance of computational difficulty will increase when larger systems need to be analyzed, and it will be become essential to pursue important breakthroughs in this regard.

10.4 Concluding Remarks

The various accomplishments of this dissertation and the above future work recommendations are directed towards the development of computational methods that aid the quantification of margins and uncertainties in engineering systems, which

[1]Damage diagnosis is basically an inverse problem. There are three steps - damage identification, localization, quantification, out of which the last step of damage quantification is identical to parameter estimation or system identification in time dependent analysis. Refer to [246, 247] for methods describing the quantification of uncertainty in each of these three stages of diagnosis.

[2]Model-based prediction when performed as a function of time, is identical to prognosis. The prediction of crack growth as a function of number of load cycles in Chapter V is an example of prognosis. The methods developed in Chapter V quantified the uncertainty in prognosis by including the various sources of uncertainty – physical variability, data uncertainty, and model uncertainty.

are essential for the purpose of decision making under uncertainty. As scientific research constantly breaks through new frontiers in engineering technology and explores undiscovered territory, the importance of quantification of margins and uncertainties in decision-making will significantly increase. Therefore, uncertainty quantification methods will also need to evolve to address increasingly challenging problems and provide a substantial foundation and framework for meaningful, risk-informed, robust decision-making.

REFERENCES

[1] J.C. Helton. Quantification of margins and uncertainties: Conceptual and computational basis. *Reliability Engineering & System Safety*, 96(9):976 – 1013, 2011.

[2] P. Dadvand, R. Rossi, and E. Oñate. An object-oriented environment for developing finite element codes for multi-disciplinary applications. *Archives of computational methods in engineering*, 17(3):253–297, 2010.

[3] E.T. Jaynes. The well-posed problem. *Foundations of Physics*, 3(4):477–492, 1973.

[4] L.E. Szabó. Objective probability-like things with and without objective indeterminism. *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics*, 38(3):626–634, 2007.

[5] R. Von Mises. *Probability, statistics, and truth*. Dover Pubns, 1981.

[6] K.R. Popper. The propensity interpretation of probability. *The British journal for the philosophy of science*, 10(37):25–42, 1959.

[7] B. De Finetti. *Theory of Probability (Two Volumes)*. John Wiley & Sons, 1974.

[8] D. Calvetti and E. Somersalo. *Introduction to Bayesian scientific computing*. New York: Springer-Verlag, 2007.

[9] P.M. Lee. *Bayesian statistics*. Arnold London, UK, 2004.

[10] T. Leonard and J.S.J. Hsu. Bayesian methods. *Cambridge Books*, 2001.

[11] D. Calvetti and E. Somersalo. *Introduction to Bayesian scientific computing: ten lectures on subjective computing*, volume 2. Springer, 2007.

[12] E. Somersalo and J. Kaipio. *Statistical and computational inverse problems*. New York: Springer-Verlag, 2004.

[13] T. Bayes and M. Price. An essay towards solving a problem in the Doctrine of Chances. by the late Rev. Mr. Bayes, FRS communicated by Mr. Price, in a letter to John Canton, AMFRS. *Philosophical Transactions (1683-1775)*, pages 370–418, 1763.

[14] S.M. Stigler. *Statistics on the table: The history of statistical concepts and methods*. Harvard Univ Pr, 2002.

[15] S.M. Stigler. Laplace's 1774 memoir on inverse probability. *Statistical Science*, 1(3):359–363, 1986.

[16] S.E. Fienberg. When did Bayesian inference become Bayesian? *Bayesian analysis*, 1(1):1–40, 2006.

[17] H. Jeffreys. *Theory of probability*. Oxford University Press, USA, 1998.

[18] J. Aldrich. R.A. Fisher on Bayes and Bayes theorem. *Bayesian Analysis*, 3(1):161–170, 2008.

[19] A.W.F. Edwards. *Likelihood*. Cambridge Univ Pr, 1984.

[20] R.A. Fisher. On an absolute criterion for fitting frequency curves. *Messenger of Mathematics*, 41(1):155–160, 1912.

[21] J. Aldrich. R.A. Fisher and the making of maximum likelihood 1912-1922. *Statistical Science*, 12(3):162–176, 1997.

[22] Y. Pawitan. *In all likelihood: statistical modelling and inference using likelihood*. Oxford University Press, USA, 2001.

[23] N.D. Singpurwalla. *Reliability and risk: a Bayesian perspective*, volume 637. Wiley, Hoboken, NJ, 2006.

[24] N.D. Singpurwalla. Betting on residual life: The caveats of conditioning. *Statistics & probability letters*, 77(12):1354–1361, 2007.

[25] F.V. Jensen. *An introduction to Bayesian networks*. New York: Springer-Verlag, 1996.

[26] D. Heckerman. A tutorial on learning with Bayesian networks. *Innovations in Bayesian Networks*, 2008.

[27] A. Haldar and S. Mahadevan. *Probability, reliability, and statistical methods in engineering design*. John Wiley & Sons, Inc., 2000.

[28] A. Haldar and S. Mahadevan. *Reliability assessment using stochastic finite element analysis*. New York: John Wiley & Sons, 2000.

[29] M. Rosenblatt. Remarks on some nonparametric estimates of a density function. *The Annals of Mathematical Statistics*, pages 832–837, 1956.

[30] C.P. Robert and G. Casella. *Monte Carlo statistical methods*. New York: Springer-Verlag, 2004.

[31] R. Rackwitz and B. Flessler. Structural reliability under combined random load sequences. *Computers & Structures*, 9(5):489–494, 1978.

[32] A. Der Kiureghian, H.Z. Lin, and S.J. Hwang. Second-order reliability approximations. *Journal of Engineering Mechanics*, 113:1208, 1987.

[33] A. Der Kiureghian and T. Dakessian. Multiple design points in first and second-order reliability. *Structural Safety*, 20(1):37–49, 1998.

[34] L. Tvedt. Distribution of quadratic forms in normal spaceapplication to structural reliability. *Journal of Engineering Mechanics*, 116:1183, 1990.

[35] A. Der Kiureghian, Y. Zhang, and C.C. Li. Inverse reliability problem. *Journal of engineering mechanics*, 120:1154, 1994.

[36] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola. *Global sensitivity analysis: The primer*. John Wiley & Sons, 2008.

[37] I.M. Sobol. Sensitivity analysis for non-linear mathematical models. *Mathematical Modeling and Computational Experiment*, 1(1):407–414, 1993.

[38] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087, 1953.

[39] R.M. Neal. Slice sampling. *Annals of Statistics*, pages 705–741, 2003.

[40] W.K. Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97, 1970.

[41] S. Geman and D. Geman. Stochastic relaxation, gibbs distributions and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.

[42] J.S. Liu, F. Liang, and W.H. Wong. The multiple-try method and local optimization in metropolis sampling. *Journal of the American Statistical Association*, pages 121–134, 2000.

[43] G.O. Roberts and J.S. Rosenthal. Harris recurrence of metropolis-within-gibbs and trans-dimensional markov chains. *The Annals of Applied Probability*, 16(4):2123–2139, 2006.

[44] M.R. Rajashekhar and B.R. Ellingwood. A new look at the response surface approach for reliability analysis. *Structural safety*, 12(3):205–220, 1993.

[45] R. Ghanem and P.D. Spanos. Polynomial chaos in stochastic finite elements. *Journal of Applied Mechanics*, 57(1):197–202, 1990.

[46] B.E. Boser, I.M. Guyon, and V.N. Vapnik. A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory*, pages 144–152. ACM, 1992.

[47] M.E. Tipping. Sparse Bayesian learning and the relevance vector machine. *The Journal of Machine Learning Research*, 1:211–244, 2001.

[48] C.E. Rasmussen. Gaussian processes in machine learning. *Advanced Lectures on Machine Learning*, pages 63–71, 2004.

[49] T.J. Santner, B.J. Williams, and W. Notz. *The design and analysis of computer experiments*. New York: Springer-Verlag, 2003.

[50] B.J. Bichon, M.S. Eldred, L.P. Swile, S. Mahadevan, and J.M. McFarland. Efficient global reliability analysis for nonlinear implicit performance functions. *AIAA journal*, 46(10):2459–2468, 2008.

[51] C.E. Rasmussen. *Evaluation of Gaussian processes and other methods for non-linear regression*. PhD thesis, University of Toronto, 1996.

[52] J.M. McFarland. *Uncertainty Analysis for Computer Simulations through Validation and Calibration*. PhD thesis, Vanderbilt University, 2008.

[53] N. Cressie. *Spatial statistics*. John Wiley and Sons, 1991.

[54] J.P. Chiles and P. Delfiner. *Geostatistics: modeling spatial uncertainty*, volume 344. Wiley-Interscience, 1999.

[55] H. Wackernagel. *Multivariate geostatistics: an introduction with applications*. New York: Springer-Verlag, 2003.

[56] V. Hombal and S. Mahadevan. Bias minimization in gaussian process surrogate modeling for uncertainty quantification. *International Journal for Uncertainty Quantification*, 1(4):321–349, 2011.

[57] W.L. Oberkampf, J.C. Helton, C.A. Joslyn, S.F. Wojtkiewicz, and S. Ferson. Challenge problems: uncertainty in system response given uncertain parameters. *Reliability Engineering & System Safety*, 85(1):11–19, 2004.

[58] X. Du, A. Sudjianto, and B. Huang. Reliability-based design with the mixture of random and interval variables. *Journal of mechanical design*, 127(6):1068–1076, 2005.

[59] S. Ferson, V. Kreinovich, J. Hajagos, W. Oberkampf, and L. Ginzburg. Experimental uncertainty estimation and statistics for data having interval uncertainty. Technical report, Sandia National Laboratories, Albuquerque, New Mexico, 2007.

[60] S. Guiasu and A. Shenitzer. The principle of maximum entropy. *The mathematical intelligencer*, 7(1):42–48, 1985.

[61] K. Zaman, S. Rangavajhala, M.P. McDonald, and S. Mahadevan. A probabilistic approach for representation of interval uncertainty. *Reliability Engineering & System Safety*, 96(1):117–130, 2011.

[62] V. Montgomery. *New statistical methods in risk assessment by probability bounds*. PhD thesis, Durham University, Durham, UK., 2009.

[63] G. Shafer. *A mathematical theory of evidence*, volume 1. Princeton university press, 1976.

[64] H.R. Bae, R.V. Grandhi, and R.A. Canfield. An approximation approach for uncertainty quantification using evidence theory. *Reliability Engineering & System Safety*, 86(3):215–225, 2004.

[65] H. Agarwal, J.E. Renaud, E.L. Preston, and D. Padmanabhan. Uncertainty quantification using evidence theory in multidisciplinary design optimization. *Reliability Engineering & System Safety*, 85(1):281–294, 2004.

[66] J.C. Helton, J.D. Johnson, and W.L. Oberkampf. An exploration of alternative approaches to the representation of uncertainty in model predictions. *Reliability Engineering & System Safety*, 85(1):39–71, 2004.

[67] Y. Ben-Haim and I. Elishakoff. *Convex models of uncertainty in applied mechanics*, volume 112. Elsevier Amsterdam, 1990.

[68] Y. Ben-Haim. A non-probabilistic measure of reliability of linear systems based on expansion of convex models. *Structural Safety*, 17(2):91–109, 1995.

[69] D. Dubois, H. Prade, and E.F. Harding. *Possibility theory: an approach to computerized processing of uncertainty*, volume 2. Plenum Press New York, 1988.

[70] S.S. Rao and K.K. Annamdas. Dempster-shafer theory in the analysis and design of uncertain engineering systems. *Product Research*, pages 135–160, 2009.

[71] G.A.F. Seber and C.J. Wild. *Nonlinear regression*. John Wiley, New York, 1989.

[72] A. Der Kiureghian. Measures of structural safety under imperfect states of knowledge. *Journal of Structural Engineering*, 115(5):1119–1140, 1989.

[73] A. Der Kiureghian. Analysis of structural reliability under parameter uncertainties. *Probabilistic engineering mechanics*, 23(4):351–358, 2008.

[74] R. Zhang and S. Mahadevan. Integration of computation and testing for reliability estimation. *Reliability Engineering & System Safety*, 74(1):13–21, 2001.

[75] O. Ditlevsen and H.O. Madsen. *Structural reliability methods*. John Wiley : New York, 1996.

[76] J.G. Hajagos. Interval monte carlo as an alternative to second-order sampling for estimating ecological risk. *Reliable computing*, 13(1):71–81, 2007.

[77] E.F. Halpern, M.C. Weinstein, M.G.M. Hunink, and G.S. Gazelle. Representing both first-and second-order uncertainties by monte carlo simulation for groups of patients. *Medical Decision Making*, 20(3):314–322, 2000.

[78] G.A. Barnard, G.M. Jenkins, and C.B. Winsten. Likelihood inference and time series. *Journal of the Royal Statistical Society. Series A (General)*, pages 321–372, 1962.

[79] W.Q. Meeker. *Statistical methods for reliability data.* John Wiley & Sons : New York, 1998.

[80] C.E. Rasmussen. The infinite gaussian mixture model. *Advances in neural information processing systems*, 12:554–560, 2000.

[81] S. Ferson and L.R. Ginzburg. Different methods are needed to propagate ignorance and variability. *Reliability Engineering & System Safety*, 54(2):133–144, 1996.

[82] K. Pearson. On the probability that two independent distributions of frequency are really samples from the same population. *Biometrika*, 8(1/2):250–254, 1911.

[83] F.J. Massey Jr. The kolmogorov-smirnov test for goodness of fit. *Journal of the American Statistical Association*, pages 68–78, 1951.

[84] T.W. Anderson and D.A. Darling. A test of goodness of fit. *Journal of the American Statistical Association*, pages 765–769, 1954.

[85] H. Cramer. *Mathematical methods of statistics*, volume 9. Princeton Univ Pr, 1999.

[86] S. Ferson, V. Kreinovich, L. Ginzburg, D.S. Myers, and K. Sentz. Constructing probability boxes and dempster-shafer structures. Technical report, Sandia National Laboratories, Albuquerque, New Mexico, 2002.

[87] D. Madigan, A.E. Raftery, C. Volinsky, and J.A. Hoeting. Bayesian model averaging. In *Proceedings of the AAAI Workshop on Integrating Multiple Learned Models, Portland, OR*, pages 77–83, 1996.

[88] J.A. Hoeting, D. Madigan, A.E. Raftery, and C.T. Volinsky. Bayesian model averaging: A tutorial. *Statistical science*, 14(4):382–401, 1999.

[89] J.O. Berger and L.R. Pericchi. The intrinsic Bayes factor for model selection and prediction. *Journal of the American Statistical Association*, 91(433):109–122, 1996.

[90] J.M. Bernardo and R. Rueda. Bayesian hypothesis testing: A reference approach. *International Statistical Review*, 70(3):351–372, 2002.

[91] R. Rebba, S. Mahadevan, and S. Huang. Validation and error estimation of computational models. *Reliability Engineering & System Safety*, 91(10):1390–1397, 2006.

[92] R. Zhang and S. Mahadevan. Bayesian methodology for reliability model acceptance. *Reliability Engineering & System Safety*, 80(1):95–103, 2003.

[93] J.H. Ahlberg, E.N. Nilson, and J.L. Walsh. *The theory of splines and their applications*, volume 38. Academic Press: New York, 1967.

[94] I.O. Kozine and L.V. Utkin. An approach to combining unreliable pieces of evidence and their propagation in a system response analysis. *Reliability Engineering & System Safety*, 85(1):103–112, 2004.

[95] G. De Cooman and M. Troffaes. Coherent lower previsions in systems modelling: products and aggregation rules. *Reliability Engineering & System Safety*, 85(1):113–134, 2004.

[96] S. Ferson and J.G. Hajagos. Arithmetic with uncertain numbers: rigorous and (often) best possible answers. *Reliability Engineering & System Safety*, 85(1):135–152, 2004.

[97] S. Ferson, C.A. Joslyn, J.C. Helton, W.L. Oberkampf, and K. Sentz. Summary from the epistemic uncertainty workshop: consensus amid diversity. *Reliability Engineering & System Safety*, 85(1):355–369, 2004.

[98] K.F. Alvin, W.L. Oberkampf, K.V. Diegert, and B.M. Rutherford. Uncertainty quantification in computational structural dynamics: A new paradigm for model validation. In *Society for Experimental Mechanics, Inc, 16th International Modal Analysis Conference.*, volume 2, pages 1191–1198. SEM, 1998.

[99] H. Akaike. A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, 19(6):716–723, 1974.

[100] H. Akaike. Likelihood and the Bayesian procedure. *Bayesian Statistics*, pages 143–166, 1980.

[101] K.P. Burnham and D.R. Anderson. *Model selection and multimodel inference: a practical information-theoretic approach*. New York: Springer-Verlag, 2002.

[102] J.L. Beck and K.V. Yuen. Model selection using response measurements: Bayesian probabilistic approach. *Journal of Engineering Mechanics*, 130:192, 2004.

[103] AIAA. Guide for the verification and validation of computational fluid dynamics simulations. Technical report, AIAA, Reston, VA, 1998.

[104] Defense Modeling and Simulation Office. Verification, validation, and accreditation (VV & A) recommended practices guide. Technical report, Office of the Director of Defense Research and Engineering, Alexandria, VA, 1998.

[105] W.L. Oberkampf and T.G. Trucano. Verification and validation in computational fluid dynamics. *Progress in Aerospace Sciences*, 38(3):209–272, 2002.

[106] P.J. Roache. Verification and validation in computational science and engineering. 1998. *Albuquerque, NM: Hermosa Publishers*, 1998.

[107] P.J. Roache. Verification of codes and calculations: Special section: Credible computational fluid dynamics simulations. *AIAA journal*, 36(5):696–702, 1998.

[108] B. Liang and S. Mahadevan. Error and uncertainty quantification and sensitivity analysis in mechanics computational models. *International Journal for Uncertainty Quantification*, 1(2):147–161, 2011.

[109] S. Rangavajhala, V. Sura, V. Hombal, and S. Mahadevan. Discretization error estimation in multidisciplinary simulations. *AIAA Journal*, 49(12):2673–2712, 2011.

[110] M. Peric and J.H. Ferziger. *Computational methods for fluid dynamics*. New York: Springer-Verlag, 1996.

[111] M. Ainsworth and J.T. Oden. A posteriori error estimation in finite element analysis. *Computer Methods in Applied Mechanics and Engineering*, 142(1-2):1–88, 1997.

[112] S.A. Richards. Completed richardson extrapolation in space and time. *Communications in numerical methods in engineering*, 13(7):573–582, 1997.

[113] W.L. Oberkampf, S.M. DeLand, B.M. Rutherford, K.V. Diegert, and K.F. Alvin. Error and uncertainty in modeling and simulation. *Reliability Engineering & System Safety*, 75(3):333–357, 2002.

[114] I. Babuška and W.C. Rheinboldt. A-posteriori error estimates for the finite element method. *International Journal for Numerical Methods in Engineering*, 12(10):1597–1615, 1978.

[115] L. Demkowicz, J.T. Oden, and T. Strouboulis. Adaptive finite elements for flow problems with moving boundaries. part i: Variational principles and a posteriori estimates. *Computer Methods in Applied Mechanics and Engineering*, 46(2):217 – 251, 1984.

[116] R.G. Ghanem and P.D. Spanos. *Stochastic finite elements: a spectral approach*. Courier Dover Pubns, 2003.

[117] M.D. Buhmann. *Radial basis functions: theory and implementations*, volume 12. Cambridge Univ Pr, 2003.

[118] C.E. Rasmussen and C.K.I. Williams. *Gaussian Processes for Machine Learning*. The MIT Press, Cambridge, MA, USA, 2006.

[119] O.M. Alifanov. Inverse heat transfer problems. *Moscow Izdatel Mashinostroenie*, 1, 1988.

[120] L. Ljung. *System identification*. Wiley Online Library, 1999.

[121] J. Hadamard. Sur les problemes aux dérivées partielles et leur signification physique. *Princeton University Bulletin*, 13(49-52):28, 1902.

[122] A. Tarantola. *Inverse problem theory and methods for model parameter estimation*. Society for Industrial Mathematics, 2005.

[123] N.R. Draper and H. Smith. *Applied regression analysis (wiley series in probability and statistics)*. Wiley-Interscience, 1998.

[124] S. Van Huffel and J. Vandewalle. *The total least squares problem: computational aspects and analysis*. Society for Industrial Mathematics, 1991.

[125] W.E. Deming. *Statistical adjustment of data*. Wiley, 1943.

[126] J. Kiefer. Sequential minimax search for a maximum. *Proc. Amer. Math. Soc*, 4(3):502–506, 1953.

[127] R.P. Brent. *Algorithms for minimization without derivatives*. Dover Pubns, 2002.

[128] P.M. Reilly and H. Patino-Leal. A Bayesian study of the error-in-variables model. *Technometrics*, pages 221–231, 1981.

[129] J.M. Marin, K. Mengersen, and C.P. Robert. Bayesian modeling and inference on mixtures of distributions. *Handbook of statistics*, 25:459–507, 2005.

[130] H. Jeffreys. An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, pages 453–461, 1946.

[131] A.N. Tikhonov and V.I.A. Arsenin. *Solutions of ill-posed problems*. Winston Washington, DC:, 1977.

[132] W.M. McKeeman. Algorithm 145: Adaptive numerical integration by simpson's rule. *Communications of the ACM*, 5(12):604, 1962.

[133] D.O. Smallwood, D.L. Gregory, and R.G. Coleman. A three parameter constitutive model for a joint which exhibits a power law relationship between energy loss and relative displacement. In *Proceedings of the 72nd Shock and Vibration Symposium*, Destin, FL, 2001.

[134] A. Urbina, T. Paez, D. Gregory, and B. Resor. Probabilistic modeling of mechanical joints. *Noise-Con, Cleveland, OH*, 2003.

[135] S. Ferson, W.L. Oberkampf, and L. Ginzburg. Model validation and predictive capability for the thermal challenge problem. *Computer Methods in Applied Mechanics and Engineering*, 197(29):2408–2430, 2008.

[136] S. Kullback and R.A. Leibler. On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86, 1951.

[137] R. Beran. Minimum hellinger distance estimates for parametric models. *The Annals of Statistics*, 5(3):445–463, 1977.

[138] A. Bhattacharyya. On a measure of divergence between two statistical populations defined by their probability distributions. *Bull. Calcutta Math. Soc*, 35(99-109):4, 1943.

[139] T.G. Trucano. Aspects of ASCI code verification and validation. Technical report, Sandia National Laboratories, Albuquerque, NM, 2000.

[140] W.L. Oberkampf and M.F. Barone. Measures of agreement between computation and experiment: validation metrics. *Journal of Computational Physics*, 217(1):5–36, 2006.

[141] R.G. Hills and I.H. Leslie. Statistical validation of engineering and scientific models: Validation experiments to application. Technical report, Sandia National Labs., Albuquerque, NM (US); Sandia National Labs., Livermore, CA (US), 2003.

[142] A. Urbina, T.L. Paez, T. Hasselman, W. Wathugala, and K. Yap. Assessment of model accuracy relative to stochastic system behavior. In *44th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, 2003.

[143] R. Rebba and S. Mahadevan. Computational methods for model reliability assessment. *Reliability Engineering & System Safety*, 93(8):1197–1207, 2008.

[144] R. Rebba. *Model validation and design under uncertainty*. PhD thesis, Vanderbilt University, 2005.

[145] S. Mahadevan and R. Rebba. Validation of reliability computational models using Bayes networks. *Reliability Engineering & System Safety*, 87(2):223–232, 2005.

[146] R. Rebba and S. Mahadevan. Validation of models with multivariate output. *Reliability Engineering & System Safety*, 91(8):861–871, 2006.

[147] X. Jiang and S. Mahadevan. Bayesian risk-based decision method for model validation under uncertainty. *Reliability Engineering & System Safety*, 92(6):707–718, 2007.

[148] W.L. Oberkampf and C.J. Roy. *Verification and validation in scientific computing*. Cambridge Univ Pr, 2010.

[149] C.J. Roy and W.L. Oberkampf. A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing. *Computer Methods in Applied Mechanics and Engineering*, 2011.

[150] R.E. Caflisch. Monte carlo and quasi-monte carlo methods. *Acta numerica*, 1998:1–49, 1998.

[151] D.O. Harris. *Chapter in Reliability-based Mechanical Design, Edited by Cruse, T.A.*, chapter Probabilistic Crack Growth. Marcel Dekker Inc., 1997.

[152] S. Rahman. Probabilistic fracture mechanics: J-estimation and finite element methods. *Engineering Fracture Mechanics*, 68(1):107–125, 2001.

[153] J.W. Provan. Probabilistic fracture mechanics and reliability. *Martinus Nijhoff Publishers, P. O. Box 163, 3300 AD Dordrecht, The Netherlands, 1987. 467*, 1987.

[154] G.H. Besterfield, W.K. Liu, M.A. Lawrence, and T. Belytschko. Fatigue crack growth reliability by probabilistic finite elements. *Computer methods in applied mechanics and engineering*, 86(3):297–320, 1991.

[155] S. Rahman and B.N. Rao. Probabilistic fracture mechanics by galerkin meshless methods–part ii: reliability analysis. *Computational mechanics*, 28(5):365–374, 2002.

[156] G.R. Johnson and W.H. Cook. Fracture characteristics of three metals subjected to various strains, strain rates, temperatures and pressures. *Engineering Fracture Mechanics*, 21(1):31–48, 1985.

[157] G. Maymon. Probabilistic crack growth behavior of aluminum 2024-t351 alloy using the unifiedapproach. *International journal of fatigue*, 27(7):828–834, 2005.

[158] S.W. Doebling and F.M. Hemez. Overview of uncertainty assessment for structural health monitoring. In *Proceedings of the 3rd International Workshop on Structural Health Monitoring*, 2001.

[159] F.M. Hemez, A.N. Robertson, and A.C. Rutherford. Uncertainty quantification and model validation for damage prognosis. In *Proceedings of the 3rd Internation Workshop on Structural Health Monitoring*, Stanford University, Stanford, California, 2003.

[160] C.R. Farrar, G. Park, F.M. Hemez, T.B. Tippetts, H. Sohn, J. Wait, and D.W. Allen. Damage detection and prediction for composite plates (u). *Journal of The Minerals, Metals and Materials Society*, 2004.

[161] C.R. Farrar and N.A.J. Lieven. Damage prognosis: the future of structural health monitoring. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1851):623–632, 2007.

[162] S.G. Pierce, K. Worden, and A. Bezazi. Uncertainty analysis of a neural network used for fatigue lifetime prediction. *Mechanical Systems and Signal Processing*, 22(6):1395–1411, 2008.

[163] G. Yagawa, Y. Kanto, S. Yoshimura, and K. Shibata. Recent research activity of probabilistic fracture mechanics for nuclear structural components in Japan. In *16th Int. Conf. SMiRT, Washington DC, USA*, 2001.

[164] R. Patrick, M.E. Orchard, B. Zhang, M.D. Koelemay, G.J. Kacprzynski, A.A. Ferri, and G.J. Vachtsevanos. An integrated approach to helicopter planetary gear fault diagnosis and failure prognosis. In *Autotestcon, 2007 IEEE*, pages 547–552. IEEE, 2007.

[165] R.G. Tryon, T.A. Cruse, and S. Mahadevan. Development of a reliability-based fatigue life model for gas turbine engine structures. *Engineering fracture mechanics*, 53(5):807–828, 1996.

[166] M. Orchard, G. Kacprzynski, K. Goebel, B. Saha, and G. Vachtsevanos. Advances in uncertainty representation and management for particle filtering applied to prognostics. In *Prognostics and Health Management, 2008. PHM 2008. International Conference on*, pages 1–6. IEEE, 2008.

[167] J.M. Papazian, E.L. Anagnostou, S.J. Engel, D. Hoitsma, J. Madsen, R.P. Silberstein, G. Welsh, and J.B. Whiteside. A structural integrity prognosis system. *Engineering Fracture Mechanics*, 76(5):620–632, 2009.

[168] R.C. McClung, M.P. Enright, H.R. Millwater, G.R. Leverant, and S.J. Hudak Jr. A software framework for probabilistic fatigue life assessment of gas turbine engine rotors. *J. ASTM International*, 1(8):199, 2004.

[169] DARWIN. *Design assessment for reliability with inspection*. Software developed at Southwest Research institute, 2002.

[170] GENOA. *Genoa 4.2*. Genow Virtual testing and Analysis Software, 2007.

[171] W.A. Grell and P.J. Laz. Probabilistic fatigue life prediction using afgrow and accounting for material variability. *International Journal of Fatigue*, 32(7):1042–1049, 2010.

[172] J.C. Newman Jr., A. Brot, and C. Matias. Crack-growth calculations in 7075-t7351 aluminum alloy under various load spectra using an improved crack-closure model. *Engineering fracture mechanics*, 71(16-17):2347–2363, 2004.

[173] D. Virkler, B.M. Hillberry, and P.K. Goel. The statistical nature of fatigue crack propagation, air force flight dynamics laboratory. Technical report, AFFDL-TR. 78-43, WPAFB, Air Force Research Laboratory, 1978.

[174] Y.K. Lin and J.N. Yang. A stochastic theory of fatigue crack propagation. *AIAA Journal(ISSN 0001-1452)*, 23:117–124, 1985.

[175] J.N. Yang and S.D. Manning. A simple second order approximation for stochastic crack growth analysis. *Engineering fracture mechanics*, 53(5):677–686, 1996.

[176] W.F. Wu and C.C. Ni. Probabilistic models of fatigue crack propagation and their experimental verification. *Probabilistic Engineering Mechanics*, 19(3):247–257, 2004.

[177] A. Saxena, J. Celaya, B. Saha, S. Saha, and K. Goebel. Evaluating algorithm performance metrics tailored for prognostics. In *Aerospace conference, 2009 IEEE*, pages 1–13. IEEE, 2009.

[178] A. Urbina. *Uncertainty Quantification and Decision Making in Hierarchical Development of Computational Models*. PhD thesis, Doctoral Thesis Dissertation. Vanderbilt University, Nashville, TN, 2009.

[179] S. Sankararaman, Y. Ling, C. Shantz, and S. Mahadevan. Uncertainty quantification in fatigue damage prognosis. In *Annual Conference of the Prognostics and Health Management Society*, 2009.

[180] A. Makeev, Y. Nikishkov, and E. Armanios. A concept for quantifying equivalent initial flaw size distribution in fracture mechanics based life prediction models. *International journal of fatigue*, 29(1):141–145, 2007.

[181] S. Sankararaman, Y. Ling, and S. Mahadevan. Statistical inference of equivalent initial flaw size with complicated structural geometry and multi-axial variable amplitude loading. *International Journal of Fatigue*, 32(10):1689–1700, 2010.

[182] R. Cross, A. Makeev, and E. Armanios. Simultaneous uncertainty quantification of fracture mechanics based life prediction model parameters. *International journal of fatigue*, 29(8):1510–1515, 2007.

[183] P.C. Paris. A rational analytic theory of fatigue. *Trends Engin*, 13:9–14, 1961.

[184] O.E. Wheeler. Spectrum loading and crack growth. *Journal of Basic Engineering*, 94(01), 1972.

[185] B.K.C. Yuen and F. Taheri. Proposed modifications to the Wheeler retardation model for multiple overloading fatigue life prediction. *International journal of fatigue*, 28(12):1803–1819, 2006.

[186] B.C. Sheu, P.S. Song, and S. Hwang. Shaping exponent in Wheeler model under a single overload. *Engineering fracture mechanics*, 51(1):135–143, 1995.

[187] P.S. Song, B.C. Sheu, and L. Chang. A modified Wheeler model to improve predictions of crack growth following a single overload. *JSME International Journal Series A*, 44(1):117–122, 2001.

[188] Y. Liu and S. Mahadevan. Probabilistic fatigue life prediction using an equivalent initial flaw size distribution. *International Journal of Fatigue*, 31(3):476–487, 2009.

[189] Y. Liu and S. Mahadevan. Multiaxial high-cycle fatigue criterion and life prediction for metals. *International Journal of Fatigue*, 27(7):790–800, 2005.

[190] Y. Liu, L. Liu, and S. Mahadevan. Analysis of subsurface crack propagation under rolling contact loading in railroad wheels using fem. *Engineering fracture mechanics*, 74(17):2659–2674, 2007.

[191] Y. Liu and S. Mahadevan. Threshold stress intensity factor and crack growth rate prediction under mixed-mode loading. *Engineering fracture mechanics*, 74(3):332–345, 2007.

[192] R. Zhang and S. Mahadevan. Fatigue reliability analysis using nondestructive inspection. *Journal of Structural Engineering*, 127:957, 2001.

[193] ANSYS. *ANSYS Theory reference*. ANSYS Inc., 2007.

[194] Y. Ling, C. Shantz, S. Mahadevan, and S. Sankararaman. Stochastic prediction of fatigue loading using real-time monitoring data. *International Journal of Fatigue*, 33(7):868 – 879, 2011.

[195] Y. Ling and S. Mahadevan. Integration of structural health monitoring and fatigue damage prognosis. *Mechanical Systems and Signal Processing*, In Press(Available online), 2011.

[196] A. Urbina, S. Mahadevan, and T.L. Paez. Quantification of margins and uncertainties of complex systems in the presence of aleatoric and epistemic uncertainty. *Reliability Engineering & System Safety*, 2011.

[197] S. Sankararaman, Y. Ling, and S. Mahadevan. Uncertainty quantification and model validation of fatigue crack growth prediction. *Engineering Fracture Mechanics*, 78(7):1487 – 1504, 2011.

[198] S. Sankararaman, Y. Ling, C. Shantz, and S. Mahadevan. Uncertainty quantification in fatigue crack growth prognosis. *International Journal of Prognosis and Health Monitoring*, 2(1):15, 2011.

[199] S. Sankararaman, Y. Ling, C. Shantz, and S. Mahadevan. Inference of equivalent initial flaw size under multiple sources of uncertainty. *International Journal of Fatigue*, 33(2):75–89, 2011.

[200] J.R. Red-Horse and T.L. Paez. Sandia national laboratories validation workshop: structural dynamics application. *Computer Methods in Applied Mechanics and Engineering*, 197(29-32):2578–2584, 2008.

[201] M.C. Kennedy and A. O'Hagan. Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3):425–464, 2001.

[202] W.R. Gilks, S. Richardson, and D.J. Spiegelhalter. *Markov chain Monte Carlo in practice*. Chapman & Hall, 1996.

[203] W.R. Gilks and P. Wild. Adaptive rejection sampling for gibbs sampling. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 41(2):337–348, 1992.

[204] A.K. Chopra. *Dynamics of structures*. Prentice Hall, 1995.

[205] E.J. Cramer, J.E. Dennis Jr, P.D. Frank, R.M. Lewis, and G.R. Shubin. Problem formulation for multidisciplinary optimization. *SIAM Journal on Optimization*, 4:754–776, 1994.

[206] N.M. Alexandrov and R.M. Lewis. Algorithmic perspectives on problem formulations in mdo. In *Proceedings of the 8th AIAA/USAF/NASA/ISSMO Symposium on MA & O, Long Beach, CA*. AIAA, 2000.

[207] T. Belytschko. Fluid-structure interaction. *Computers & Structures*, 12(4):459–469, 1980.

[208] E.A. Thornton. *Thermal structures for aerospace applications*. AIAA, 1996.

[209] A.J. Culler, A.R. Crowell, and J.J. McNamara. Studies on fluid-structural coupling for aerothermoelasticity in hypersonic flow. *AIAA journal*, 48(May):1721–1738, 2009.

[210] C.A. Felippa, K.C. Park, and C. Farhat. Partitioned analysis of coupled mechanical systems. *Computer methods in applied mechanics and engineering*, 190(24-25):3247–3270, 2001.

[211] C. Michler, S.J. Hulshoff, E.H. Van Brummelen, and R. De Borst. A monolithic approach to fluid–structure interaction. *Computers & fluids*, 33(5):839–848, 2004.

[212] K.C. Park, C.A. Felippa, and J.A. DeRuntz. Stabilization of staggered solution procedures for fluid-structure interaction analysis. *Computational methods for fluid-structure interaction problems*, pages 95–124, 1977.

[213] R.D. Braun. *Collaborative optimization: an architecture for large-scale distributed design*. PhD thesis, Stanford University, 1996.

[214] R.D. Braun, A.A. Moore, and I.M. Kroo. Collaborative approach to launch vehicle design. *Journal of Spacecraft and Rockets*, 34(4):478–486, 1997.

[215] J. Sobieszczanski-Sobieski. Optimization by decomposition: A step from hierarchic to non-hierarchic systems. *Recent Advances in Multidisciplinary Analysis and Optimization*, 1:51–78, 1988.

[216] B.A. Wujek, J.E. Renaud, and S.M. Batill. A concurrent engineering approach for multidisciplinary design in a distributed computing environment. In *Multidisciplinary Design Optimization: State-of-the-Art, Proceedings in Applied Mathematics*, volume 80, pages 189–208, 1997.

[217] J. Sobieszczanski-Sobieski, T.D. Altus, M. Phillips, and R. Sandusky. Bilevel integrated system synthesis for concurrent and distributed processing. *AIAA journal*, 41(10):1996–2003, 2003.

[218] M. Kokkolaras, Z.P. Mourelatos, and P.Y. Papalambros. Design optimization of hierarchically decomposed multilevel systems under uncertainty. *Journal of Mechanical Design*, 128:503, 2006.

[219] H. Liu, W. Chen, M. Kokkolaras, P.Y. Papalambros, and H.M. Kim. Probabilistic analytical target cascading: a moment matching formulation for multilevel optimization under uncertainty. *Journal of Mechanical Design*, 128:991, 2006.

[220] X. Gu, J.E. Renaud, S.M. Batill, R.M. Brach, and A.S. Budhiraja. Worst case propagated uncertainty of multidisciplinary systems in robust design optimization. *Structural and Multidisciplinary Optimization*, 20(3):190–213, 2000.

[221] X. Du and W. Chen. Collaborative reliability analysis under the framework of multidisciplinary systems design. *Optimization and Engineering*, 6(1):63–84, 2005.

[222] S. Mahadevan and N. Smith. Efficient first-order reliability analysis of multidisciplinary systems. *International Journal of Reliability and Safety*, 1(1):137–154, 2006.

[223] X. Zhang and H.Z. Huang. Sequential optimization and reliability assessment for multidisciplinary design optimization under aleatory and epistemic uncertainties. *Structural and Multidisciplinary Optimization*, 40(1):165–175, 2010.

[224] M. Li and S. Azarm. Multiobjective collaborative robust optimization with interval uncertainty and interdisciplinary uncertainty propagation. *Journal of Mechanical Design*, 130:081402, 2008.

[225] X. Du, J. Guo, and H. Beeram. Sequential optimization and reliability assessment for multidisciplinary systems design. *Structural and Multidisciplinary Optimization*, 35(2):117–130, 2008.

[226] A. Chiralaksanakul and S. Mahadevan. Decoupled approach to multidisciplinary design optimization under uncertainty. *Optimization and Engineering*, 8(1):21–42, 2007.

[227] N. Smith. *Probabilistic Design of Multidisciplinary Systems*. PhD thesis, Vanderbilt University, Nashville, TN, USA, 2007.

[228] K.B. Zaman. *Modeling and Management of Epistemic Uncertainty for Multidisciplinary System Analysis and Design*. PhD thesis, Vanderbilt University, Nashville, TN, USA, 2010.

[229] P. Hajela, C.L. Bloebaum, and J. Sobieszczanski-Sobieski. Application of global sensitivity equations in multidisciplinary aircraft synthesis. *Journal of Aircraft;(USA)*, 27, 1990.

[230] J.R. Wertz and W.J. Larson. *Space mission analysis and design*. Torrance, CA (United States); Microcosm, Inc., 1999.

[231] S. Ferson, W. Tucker, C. Paredis, Y. Bulgak, and V. Kreinovich. Accounting for epistemic and aleatory uncertainty in early system design. Technical report, NASA SBIR, 2009.

[232] J.C. Helton. Quantification of margins and uncertainties: Conceptual and computational basis. *Reliability Engineering & System Safety*, 96(9):976 – 1013, 2011.

[233] A. Urbina, S. Mahadevan, and T. Paez. Resource allocation using quantification of margins and uncertainty. In *51st AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, 2011.

[234] I.M. Sobol. Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. *Mathematics and Computers in Simulation*, 55(1-3):271–280, 2001.

[235] B. Sudret. Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering & System Safety*, 93(7):964–979, 2008.

[236] S. Sankararaman and S. Mahadevan. Likelihood-based representation of epistemic uncertainty due to sparse point data and/or interval data. *Reliability Engineering & System Safety*, 96(7):814 – 824, 2011.

[237] H.C. Edwards. Sierra framework for parallel adaptive multiphysics applications. Technical report, Sandia National Laboratories, 2004.

[238] M.S. Eldred, A.A. Giunta, B.G. van Bloemen Waanders, S.F. Wojtkiewicz Jr, W.E. Hart, and M. Alleva. Dakota, a multilevel parallel object-oriented framework for design optimization, parameter estimation, uncertainty quantification, and sensitivity analysis version 3.0 reference manual. Technical report, Sandia National Labs., Albuquerque, NM (US) Sandia National Labs., Livermore, CA (US), 2001.

[239] X. Jiang and S. Mahadevan. Bayesian cross-entropy methodology for optimal design of validation experiments. *Measurement Science and Technology*, 17:1895, 2006.

[240] M. Evans and T. Swartz. Methods for approximating integrals in statistics with special emphasis on Bayesian integration problems. *Statistical Science*, pages 254–272, 1995.

[241] S.W. Raudenbush, M.L. Yang, and M. Yosef. Maximum likelihood for generalized linear models with nested random effects via high-order, multivariate laplace approximation. *Journal of Computational and Graphical Statistics*, pages 141–157, 2000.

[242] N. Pugno, M. Ciavarella, P. Cornetti, and A. Carpinteri. A generalized paris' law for fatigue crack growth. *Journal of the Mechanics and Physics of Solids*, 54(7):1333–1349, 2006.

[243] J.C. Newman. *FASTRAN II: a fatigue crack growth structural analysis program*. National Aeronautics and Space Administration, Langley Research Center, 1992.

[244] J.A. Harter. AFGROW users guide and technical manual, 1999.

[245] P.M.G.P. Moreira, P.F.P. de Matos, and P.M.S.T. de Castro. Fatigue striation spacing and equivalent initial flaw size in al 2024-t3 riveted specimens. *Theoretical and applied fracture mechanics*, 43(1):89–99, 2005.

[246] S. Sankararaman and S. Mahadevan. Uncertainty quantification in structural damage diagnosis. *Structural Control and Health Monitoring*, 18(8):807–824, 2011.

[247] S. Sankararaman and S. Mahadevan. Bayesian methodology for diagnosis uncertainty quantification and health monitoring. *Structural Control and Health Monitoring*, Accepted, In Press, 2011.