

# Uncertainty Assessment of Complex Models with Application to Aviation Environmental Systems

by

Douglas Lawrence Allaire

Submitted to the Department of Aeronautics and Astronautics  
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2009

© Massachusetts Institute of Technology 2009. All rights reserved.



**ARCHIVES**

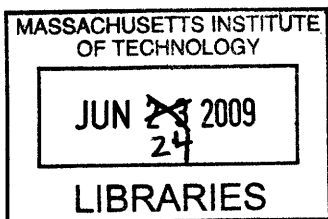
Author .....  
Department of Aeronautics and Astronautics  
May 22, 2009

Certified by .....  
Karen E. Willcox  
Associate Professor, Department of Aeronautics and Astronautics  
Thesis Supervisor

Certified by .....  
Ian A. Waitz  
Jerome C. Hunsaker Professor of Aeronautics and Astronautics  
Committee Member

Certified by .....  
Dan Frey  
Associate Professor, Department of Mechanical Engineering  
Committee Member

Accepted by .....  
David L. Darmofal  
Associate Department Head  
Chairman, Committee on Graduate Students





# Uncertainty Assessment of Complex Models with Application to Aviation Environmental Systems

by

Douglas Lawrence Allaire

Submitted to the Department of Aeronautics and Astronautics  
on May 22, 2009, in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy

## Abstract

Numerical simulation models that support decision-making and policy-making processes are often complex, involving many disciplines, and long computation times. These models typically have many factors of different character, such as operational, design-based, technological, and economics-based. Such factors generally contain uncertainty, which leads to uncertainty in model outputs. For such models, it is critical to both the application of model results and the future development of the model that uncertainty be properly assessed.

This thesis presents a comprehensive approach to the uncertainty assessment of complex models intended to support decision- and policy-making processes. The approach consists of seven steps, which are establishing assessment goals, documenting assumptions and limitations, documenting model factors and outputs, classifying and characterizing factor uncertainty, conducting uncertainty analysis, conducting sensitivity analysis, and presenting results. Factor uncertainty is represented probabilistically, characterized by the principle of maximum uncertainty, and propagated via Monte Carlo simulation. State-of-the-art methods of global sensitivity analysis are employed to apportion model output variance across model factors, and a fundamental extension of global sensitivity analysis, termed distributional sensitivity analysis, is developed to determine on which factors future research should focus to reduce output variability. The complete approach is demonstrated on a real-world model intended to estimate the impacts of aviation on climate change in support of decision- and policy-making, where it is established that a systematic approach to uncertainty assessment is critical to the proper application and future development of complex models.

A novel surrogate modeling methodology designed specifically for uncertainty assessment is also presented and demonstrated for an aircraft emissions prediction model that is being developed and applied to support aviation environmental policy-making. The results demonstrate how confidence intervals on surrogate model predictions can be used to balance the tradeoff between computation time and uncertainty in the estimation of statistical outputs of interest in uncertainty assessment.

Thesis Supervisor: Karen E. Willcox

Title: Associate Professor, Department of Aeronautics and Astronautics

Committee Member: Ian A. Waitz

Title: Jerome C. Hunsaker Professor of Aeronautics and Astronautics

Committee Member: Dan Frey

Title: Associate Professor, Department of Mechanical Engineering

# Acknowledgments

The work presented in this thesis represents the technical side of obtaining a Ph.D., for which I owe a great deal of thanks to several people. First, I would like to express my deepest gratitude to my advisor, Prof. Karen Willcox. I have learned so much from working with her over the past several years. I would also like to thank Prof. Ian Waitz and Prof. Dan Frey, who as members of my thesis committee provided me with valuable insights and technical support along the way, as well as my thesis readers, Dr. Laura Swiler and Prof. Youssef Marzouk.

Throughout this research I worked closely with several people from the Volpe Transportation Center that I would like to acknowledge. Those people are George Noel, Chris Roof, Andrew Hansen, and Gregg Fleming. All of the work presented in this thesis regarding the aviation environmental design tool was made possible by their assistance. I would also like to thank Becky Cointin and Maryalice Locke for their guidance and leadership of the assessment program on which my research focused.

Of course, obtaining a Ph.D. requires more than just support and guidance from excellent technical researchers. It also requires support and guidance from family and friends. There are so many people that have helped me in so many ways on this journey. On the school side, I'd like to thank my fellow researchers in the PARTNER research group and the ACDL. In particular, I'd like to thank Anuja Mahashabde and Rhea Liem for putting up with me during the final few months of thesis preparation. On the Boston side, I'd like to acknowledge "Camp Hemenway," who made it easy to step away from the stress from time-to-time and relax. I'd also like to thank my roommates Scott Mahar and Jon Jackson, as well as Cory McLean, for always being there whenever I needed anything. And of course I'd like to thank Ari, who could always make me smile and convince me that everything would work out anytime I was having trouble with anything.

Then there is Maine. Anytime on this journey things got rough, I could always go back to Maine and regroup. There I'd find my friends, Ryan Walker, Keith Legere,

Keith Morin, Joe Seigny, and many others, always happy to see me and always convinced that I would succeed. There I'd also find my family, to whom I cannot possibly fully express my gratitude for all that they have done for me. Mom, Dad, Dan, and Diana, thank you for your unending support and love. Without that, none of this would have been possible.

# Contents

<b>1</b>	<b>Introduction</b>	<b>17</b>
1.1	Motivation for Uncertainty Assessment . . . . .	18
1.2	Terminology and Scope . . . . .	18
1.3	Requirements for Uncertainty Assessment . . . . .	20
1.4	Current Practices in Uncertainty Assessment . . . . .	22
1.5	State of the Art Uncertainty Assessment Approaches . . . . .	25
1.6	Uncertainty Assessment for Complex Models . . . . .	25
1.7	Research Objectives . . . . .	26
1.8	Thesis Outline . . . . .	29
<b>2</b>	<b>A Probabilistic Approach to Uncertainty Assessment</b>	<b>31</b>
2.1	Proposed Approach to Uncertainty Assessment . . . . .	31
2.1.1	Approach Step 1: Establish the desired outcomes of the uncertainty assessment . . . . .	31
2.1.2	Approach Step 2: Documenting assumptions and limitations of the model . . . . .	33
2.1.3	Approach Step 3: Documenting factors and outputs of the model	34
2.1.4	Approach Step 4: Classify and characterize uncertainty . . . . .	35
2.1.5	Approach Step 5: Conduct uncertainty analysis . . . . .	36
2.1.6	Approach Step 6: Conduct sensitivity analysis . . . . .	36
2.1.7	Approach Step 7: Present results . . . . .	38
2.2	Characterizing Uncertainty . . . . .	38
2.3	Uncertainty Analysis . . . . .	46

<b>3</b>	<b>Sensitivity Analysis for Complex Models</b>	<b>49</b>
3.1	Sensitivity Analysis Background . . . . .	50
3.2	Global Sensitivity Analysis . . . . .	51
3.3	Distributional Sensitivity Analysis . . . . .	56
3.3.1	Derivation of distributional sensitivity analysis . . . . .	57
3.3.2	Defining reasonable distributions . . . . .	59
3.3.3	Acceptance/rejection sampling . . . . .	65
3.3.4	Example of global and distributional sensitivity analysis . . .	69
3.4	Convergence of Uncertainty and Sensitivity Analyses . . . . .	72
<b>4</b>	<b>Surrogate Modeling for Uncertainty Assessment</b>	<b>77</b>
4.1	Surrogate Modeling Background . . . . .	78
4.2	Application . . . . .	79
4.2.1	Aircraft emissions model . . . . .	79
4.3	Surrogate Modeling Methodology with Quantified Confidence Intervals	82
4.3.1	AEM surrogate for decision-making uncertainty analysis . . .	82
4.3.2	AEM surrogate for model development sensitivity analysis . .	86
4.4	Results . . . . .	90
4.4.1	AEM surrogate results for decision-making uncertainty analysis	90
4.4.2	AEM surrogate results for model development sensitivity analysis	92
<b>5</b>	<b>Uncertainty Assessment of a Real-World Aviation Environmental Model</b>	<b>97</b>
5.1	APMT-Impacts Climate Module . . . . .	97
5.1.1	Motivation for the APMT-Impacts climate module . . . . .	98
5.1.2	Modeling methodology of the climate module . . . . .	98
5.1.3	The nominal climate model . . . . .	103
5.2	Uncertainty Assessment of the APMT-Impacts Nominal Climate Model	105
5.2.1	Step 1: Establishing uncertainty assessment objectives . . . .	105
5.2.2	Step 2: Document assumptions and limitations of the model .	106
5.2.3	Step 3: Document factors and outputs . . . . .	114



5.2.4	Step 4: Classify and characterize uncertainty . . . . .	114
5.2.5	Step 5: Conduct uncertainty analysis . . . . .	114
5.2.6	Step 6: Conduct sensitivity analysis . . . . .	118
5.3	APMT-Impacts Nominal Climate Model Uncertainty Assessment Con- clusions . . . . .	121
5.4	Further Uncertainty Assessment of the Climate Module . . . . .	122
<b>6</b>	<b>Conclusions and Future Work</b>	<b>127</b>
6.1	Summary . . . . .	127
6.2	Conclusions . . . . .	128
6.3	Future Work . . . . .	129



# List of Figures

1-1	Diagram of a model with its factors, parameters, and outputs identified.	19
1-2	Diagram of the modeling process which shows the connections between the various steps. . . . .	20
1-3	FAA tools-suite for estimating the impacts of aviation on the environment [1] . . . . .	27
2-1	Example of a uniform probability density function. . . . .	41
2-2	Example of a discrete uniform probability mass function. . . . .	42
2-3	Example of a normal probability density function. . . . .	43
2-4	Example of a beta probability density function. . . . .	44
2-5	Example of a triangular density function. . . . .	45
3-1	Apportioning Output Variance . . . . .	52
3-2	Example of a uniform probability density function on [a,b]. . . . .	60
3-3	Example of a triangular probability density function with minimum value = a, maximum value = b, and mode = c. . . . .	62
3-4	The top two plots are the samples drawn uniformly from the possible values of the minimum (left) and the resulting values of the maximum (right). The bottom two plots are the samples drawn using the proposed procedure for sampling from a family of triangular distributions.	64
3-5	Example of a normal probability density function with mean = $\mu$ and variance = $\sigma^2$ . . . . .	65

3-6	Example histograms that result from randomly sampling from an original uniform random variable on $[0,1]$ (top left), and the use of acceptance/rejection sampling on the original samples with $\delta = 0.5$ (top right, lower left, lower right). . . . .	67
3-7	Example histograms that result from randomly sampling from an original triangular random variable with parameters $(0,2,1)$ (top left), and the use of acceptance/rejection sampling on the original samples with $\delta = 0.5$ (top right, lower left, lower right). . . . .	67
3-8	Example histograms that result from randomly sampling from an original normal random variable with mean = 0 and variance = 1 (left), and the use of acceptance/rejection sampling on the original samples with $\delta = 0.5$ (right). . . . .	68
3-9	Adjusted main effect sensitivity indices of each factor for values of $\delta = 0.0, 0.1, 0.2, \dots, 1.0$ . The adjusted main effect sensitivity indices of each factor for $\delta = 0$ (rightmost points), are the main effect sensitivity indices of each factor as computed by global sensitivity analysis. . . .	70
3-10	A comparison of the main effect sensitivity indices estimated from global sensitivity analysis to the average adjusted main effect sensitivity indices estimated from distributional sensitivity analysis. . . .	71
3-11	Estimates of the mean and variance of the test model output as the number of iterations increases. . . . .	72
3-12	Estimates of the total, main, and average adjusted main effect sensitivity indices of factor $X_1$ as the number of iterations used for global sensitivity analysis increases. . . . .	73
3-13	Estimates of the adjusted main effect sensitivity indices for $\delta = 0.9$ (top), $\delta = 0.6$ (middle), and $\delta = 0.3$ (bottom), for factor $X_1$ as the number of iterations used for global sensitivity analysis increases. . .	74
3-14	Bootstrap confidence intervals for the adjusted main effect sensitivity indices of factor $X_1$ using 50 reasonable distributions for each value of $\delta$ . . .	74

4-1	AEM segmentation of an aircraft operation. Adapted from Ref. [2]. . . . .	80
4-2	The hierarchical surrogate modeling approach achieves a reduction in computational complexity through a reduction of the input space. For the AEM, this amounts to selecting a subset of $r$ operations, denoted by the subscripts, $i_1, \dots, i_r$ , over which to estimate the total emissions. . . . .	82
4-3	Estimates of $\mu_{\mathbb{E}}, \sigma_{\mathbb{E}}^2, \mu_{\sigma^2}$ , and $\sigma_{\sigma^2}^2$ as the number of operations in the surrogate model, $n_o$ , increases from 2,500 to 9,914. . . . .	92
4-4	95 percent confidence interval widths, in terms of percentage $\pm$ of the estimated value, for the mean and variance of total $\text{NO}_x$ emissions for a full run of the representative day as $n_o$ increases from 2,500 to 9,914. . . . .	93
4-5	Total and main effect sensitivity indices for the AEM $\text{NO}_x$ output as evaluated using a surrogate model consisting of 5,000 operations. The error bars show the 95% confidence intervals for each index. . . . .	93
4-6	95 percent confidence interval widths, in terms of percentage $\pm$ of the surrogate model sensitivity index estimates, for the total and main effect sensitivity index of the temperature input as $n_o$ increases from 2500 to 9914. . . . .	94
5-1	Overview of the modeling approach for the APMT-Impacts climate module (adapted from [3]). . . . .	99
5-2	Tornado chart of the NPV output of the climate module, which shows the impacts of epistemic modeling uncertainty. . . . .	104
5-3	Histogram of Monte Carlo simulation results for NPV in billions of 2005 US dollars. . . . .	117
5-4	Main effect sensitivity indices of the NPV output of the nominal climate model. . . . .	119
5-5	Total effect sensitivity indices of the NPV output of the nominal climate model. . . . .	120
5-6	Adjusted main effect sensitivity indices of the NPV output of the nominal climate model. . . . .	121

5-7	Main effect sensitivity indices for the temperature change due to CO <sub>2</sub> emissions from aviation after 5 years. . . . .	123
5-8	Adjusted main effect sensitivity indices for the temperature change due to CO <sub>2</sub> emissions from aviation after 5 years. . . . .	124
5-9	Averaged adjusted main effect and main effect sensitivity indices for the temperature change due to CO <sub>2</sub> emissions from aviation after 5 years. . . . .	125

# List of Tables

2.1	Maximum Entropy Probability Distributions . . . . .	40
4.1	AEM factors and their probability density functions. All factors are applied as multipliers to nominal factor values. . . . .	80
4.2	95% confidence intervals of the mean and variance of total NO <sub>x</sub> emissions computed with surrogate models of 2,500, 5,000, 7,500, and 9,914 operations. . . . .	91
4.3	95 percent confidence intervals of the total and main effect sensitivity indices for each factor of the AEM for a full run of the representative day computed with a surrogate model of 5,000 operations. . . . .	95
5.1	Nominal climate model factors and uncertainty information. . . . .	115
5.2	Nominal climate model output information for uncertainty assessment.	115
5.3	Uncertainty classification and characterization of nominal climate model factors. . . . .	116
5.4	Estimates of the mean and standard deviation of the Nominal climate model NPV for a years worth of emissions from aviation. . . . .	116
5.5	Estimates of total and main effect sensitivity indices for the NPV output of the nominal climate model. . . . .	118
5.6	Estimates of averaged adjusted main effect sensitivity indices for the NPV output of the nominal climate model. . . . .	121





# Chapter 1

## Introduction

Numerical simulation is becoming increasingly widespread as a means to support decision-making and policy-making processes. Simulation models for such applications are often complex, involving many disciplines and long computation times. These models typically have many factors of different character, such as operational, design-based, technological, and economics-based. Such factors generally contain uncertainty, which leads to uncertainty in model outputs. For complex models intended to support decision-making and policy-making processes, *uncertainty assessment*, which consists of the proper representation, characterization, and analysis of uncertainty, is critical to both model development and the application of model results.

Motivation for uncertainty assessment is given in Section 1.1, followed by the definitions of certain terms used through this work and a statement regarding the scope of uncertainty assessment in Section 1.2. The requirements of a proper uncertainty assessment are laid out in Section 1.3, which is then followed by background on current practices in uncertainty assessment in Sections 1.4 and 1.5. The objectives of this research are stated in Section 1.7, and an outline of the remainder of the thesis is given in Section 1.8.

## 1.1 Motivation for Uncertainty Assessment

The growing use of numerical simulation models in science and engineering, and the presence of uncertainty in all aspects of modeling, has naturally led to questions such as: What confidence does one have in model results? What can be done to improve confidence in model results? What are the limits in terms of applicability of model results [4, 5]? *Uncertainty analysis*, which can be defined as the determination of the uncertainty in model results that derives from uncertainty in model factors [6], and *sensitivity analysis*, which can be defined as the study of how uncertainty in the output of a model can be apportioned to different sources of uncertainty in model factors [5], provide the answers to such questions [4]. The process of conducting both uncertainty and sensitivity analyses is referred to as *uncertainty assessment*.

## 1.2 Terminology and Scope

Prior to discussing what is required of an uncertainty assessment for complex models intended to support decision- and policy-making processes, and what exists for such a task in the literature today, it is necessary to first establish the terminology used throughout this work, and the scope of uncertainty assessment in regards to the modeling process.

The need for concrete definitions stems from the fact that terms such as *input*, *variable*, *factor*, and *parameter* are occasionally used interchangeably in the literature [7, 8, 9]. Further, terms such as *model* and *output* are rarely defined. Thus, to avoid any potential confusion, the following definitions are employed here:

**Definition 1.** Model: *A specific set of parameters and operations used in the representation of a system developed for the purpose of studying that system.*

**Definition 2.** Parameter: *A quantity that determines the characteristics of a model.*

**Definition 3.** Factor: *An external input to a model that is not contained in the definition of the model itself (i.e. not a parameter).*

**Definition 4.** Output: *A model result of interest.*

The definitions of the terms *factor* and *parameter* are inspired by Ref. [10]. The definition of the term *model* is taken from Ref. [11] with some modification. To illustrate the use of these terms, consider the following simple relation and Figure 1-1:

$$f(\mathbf{x}; \alpha, \beta) = \alpha x_1 + \beta x_2. \quad (1.1)$$

Here, the factors are  $x_1$  and  $x_2$  and are represented by  $\mathbf{x} = (x_1, x_2)$ , the parameters are  $\alpha$  and  $\beta$ , the model is  $f(\cdot; \alpha, \beta)$ , and the output is  $f(\mathbf{x}; \alpha, \beta)$ . In this model, the factors  $x_1$  and  $x_2$  may take on different values, however, the parameters  $\alpha$  and  $\beta$  are part of the model definition, and thus are considered fixed. If two different parameters were used instead, say  $\gamma$  and  $\lambda$ , then according to Definition 1 we would have a new model,  $f(\cdot; \gamma, \lambda)$ .

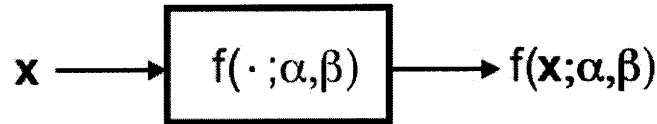


Figure 1-1: Diagram of a model with its factors, parameters, and outputs identified.

The process of developing and applying a model of a real-world system is challenging, and uncertainty assessment is an important part of that process [5, 11]. A flow chart of a general modeling process is shown in Figure 1-2, which is an adaptation from Refs. [12] and [13]. Each block in the figure represents a step in the modeling process. Model building is the process of constructing mathematical representations of the real-world phenomena in the system being studied, model verification is a process of confirming that a model is a faithful representation of what was intended, and model validation is the process of confirming the model is an adequate representation of the system and is capable of imitating the system's behavior reasonably accurately within the domain of the intended application of the model [13]. In the figure, uncertainty assessment is conducted after a model has gone through the process of model validation and verification to emphasize the role of uncertainty assessment in the ap-

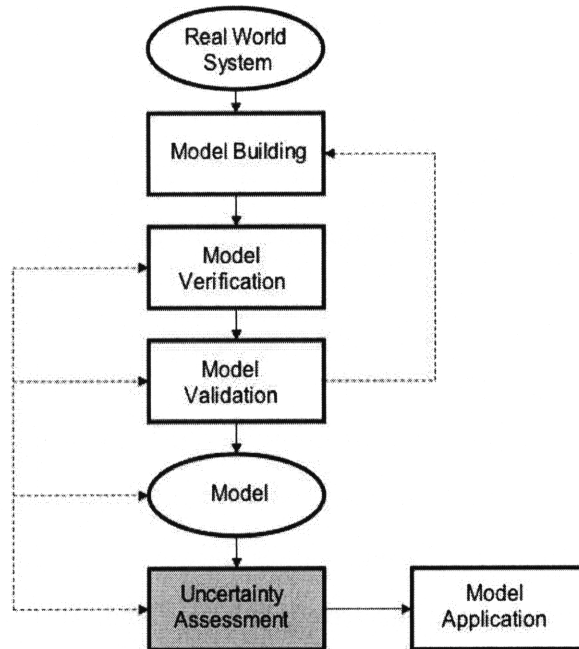


Figure 1-2: Diagram of the modeling process which shows the connections between the various steps.

plication phase of the model, where the focus is generally on uncertainty in model factors and not on uncertainty associated with the model itself. However, as shown by the dashed arrows, the modeling process is not typically conducted in series; rather steps such as verification, validation, and uncertainty assessment are complementary and oftentimes overlap. For example, uncertainty assessment can identify gaps in model functionality, which can be considered part of the verification process and can also provide model output information that can be used in the validation process. It should be noted here that the uncertainties considered in this work do not include model form uncertainty, that is uncertainty surrounding the choice of model type (e.g. the choice of an inelastic or elastic model for stress in a beam), which can at times be considerably larger than uncertainty associated with model factors considered here.

### 1.3 Requirements for Uncertainty Assessment

There are many techniques available for performing both uncertainty and sensitivity analyses, such as Monte Carlo methods [5, 14, 15], differential analysis [4, 16], and

variance-based approaches [5, 17, 18, 19]. Further, there are many different methods of representing uncertainty, such as probability theory, possibility theory, Dempster-Shafer evidence theory, imprecise probabilities, interval analysis, and several others [20, 21, 22, 23, 24]. However, this work focuses only on sampling-based probabilistic approaches to uncertainty assessment due to their general applicability, effectiveness, and wide use [6, 7, 8, 9, 20, 25].

Depending on the objectives of an uncertainty assessment (e.g. studying the sensitivity of model outputs in local regions of interest, determining which factors are responsible for most of the output variability, etc.), certain techniques may be more relevant than others. Further, prior to engaging in an uncertainty or sensitivity analysis, it is necessary to establish the types of uncertainties present and how they should be characterized, which requires careful consideration of model factors and model outputs. Finally, once uncertainty and sensitivity analyses have been carried out, results of the analyses must be presented in a meaningful manner. Thus, a formal assessment of uncertainty should include the following steps:

**Step 1:** establish the objectives of the uncertainty assessment,

**Step 2:** document assumptions and limitations of the model,

**Step 3:** document factors and outputs of the model,

**Step 4:** classify and characterize factor uncertainty,

**Step 5:** conduct uncertainty analysis,

**Step 6:** conduct sensitivity analysis,

**Step 7:** present results.

Current practices for each of these steps are described in the following section.

## 1.4 Current Practices in Uncertainty Assessment

In general, there is a wide body of literature around the various techniques of both uncertainty and sensitivity analyses, which are discussed in more detail in Chapters 2 and 3. There are also complete programs for uncertainty assessment in place such as Refs. [6, 26, 27], which will be discussed in more detail in Section 1.5, after current practices for each component of the process of uncertainty assessment have been addressed in this section. In what follows, the discussion of current practices focuses on the application of uncertainty assessment to general numerical simulation models (rather than the specific case of complex models intended to support decision- and policy-making processes), since most uncertainty assessment practices do not specify the nature of the underlying model being studied [5, 6]. The relation of these practices to an approach designed explicitly for complex models is discussed in Section 1.6.

### **Establishing objectives for uncertainty assessment**

There are many different reasons to undertake an uncertainty assessment of a model. Among these are uncovering technical errors, identifying important regions in the space of the factors, establishing research priorities, and the general defense of model results in the face of criticism [5]. Given that there are many techniques available for uncertainty and sensitivity analyses, and that these techniques may be better suited to meeting some objectives over others, it is well-established in current practice that the precise definition of objectives is a necessary first step in any approach to uncertainty assessment [4, 5].

### **Documenting assumptions and limitations of a model**

According to Ref. [4], two of the key questions that should be answered by uncertainty assessment are: How far can the calculated results of a model be extrapolated? How can the predictability and/or extrapolation limits be extended? Answering these questions requires clear identification of assumptions in place in the model and any known limitations. Generally, uncertainty assessments do not include such documentation, since the focus is typically on performing uncertainty and sensitivity analyses, rather than on qualitative aspects of model applicability [7, 9].

## Documenting factors and outputs of a model

In most uncertainty assessments, a study begins with the definition of factor distributions, which are then used to carry out uncertainty and sensitivity analyses [6, 7, 9, 28]. However, most of these assessments do not take account of the factors on which each output depends. Nevertheless, this step is important in a proper uncertainty assessment since model factors may have different types of uncertainty associated with them. These different types of factor uncertainties influence the type of uncertainty associated with model outputs, which can impact uncertainty assessment objectives, such as establishing research priorities. Further, outputs of one model may be factors in other models, and therefore, the uncertainty associated with these outputs must be clearly defined for the uncertainty assessment of those other models.

## Classifying and characterizing factor uncertainty for uncertainty assessment

It is well-established that there are different types of uncertainty that can be associated with model factors [6, 28, 29], however, in most applications of uncertainty assessment, there is a disconnect between theoretical uncertainty characterization and how uncertainty is analyzed in practice.

Uncertainty is generally classified as *aleatory*, which arises through natural randomness, or *epistemic*, which arises through imperfect knowledge [29]. The fundamental difference between these two categories is the fact that aleatory uncertainty is irreducible, whereas epistemic uncertainty may be reducible if more knowledge of the uncertainty is obtained. It is important to note factor uncertainty will usually not be neatly classed as either aleatory or epistemic. In many situations, for example atmospheric factors, there is known to be underlying aleatory uncertainty, but due to sparse data, that aleatory uncertainty cannot be precisely quantified, and thus there is a layer of epistemic uncertainty also present. For factors with this type of associated uncertainty, it is understood that some of the uncertainty is reducible, and some of it is not. In this work, any factor with some amount of uncertainty due to imperfect knowledge is classified as an epistemic uncertainty.

For uncertainty assessments, a common objective is the identification of high-

priority factors for future research, where the factors considered to be high-priority are those that are found to be responsible for most output variability [5]. However, application of state-of-the-art techniques of variance apportionment, known as global sensitivity analyses, are typically applied under the assumption that all factor uncertainty is epistemic [25, 30], which can lead to misleading conclusions. These methods are discussed in detail in Chapter 3.

The characterization of uncertainty (e.g. assigning the characteristics of a normal distribution to a particular factor’s uncertainty), is typically done inconsistently across model factors [7, 8, 9], though well-defined methods, such as the principle of maximum uncertainty exist [20]. For a proper uncertainty assessment, factor uncertainty should be characterized in a manner that is consistent, meaningful, and defensible. The characterization should be *consistent*, in that the same rules have been enforced in all uncertainty characterizations for a particular analysis; *meaningful*, in the sense that the uncertainty characterizations allow for clear interpretation of results; and *defensible*, in the sense that concrete reasons can be supplied for all decisions regarding characterization of uncertainty.

## **Uncertainty and sensitivity analyses**

As mentioned previously, uncertainty and sensitivity analyses have been applied in many different ways for many different objectives. Current best practices for these analyses are related to a given set of uncertainty assessment objectives, and thus will be discussed in Sections 2.3 and 3.1, after a set of uncertainty assessment objectives for complex models intended to support decision- and policy-making processes have been defined.

## **Presenting uncertainty assessment results**

The visual presentation of quantitative information, such as the results of an uncertainty assessment, is a distinct area of research that is usually not considered explicitly in uncertainty assessment. Most uncertainty assessment practices [5, 6] adhere to the recommendations found in [31, 32].



## 1.5 State of the Art Uncertainty Assessment Approaches

It is not typical in the practice of uncertainty assessment to cite any particular established approach; rather, practitioners usually apply various techniques of uncertainty assessment in an ad hoc fashion [7, 8, 9, 25]. Nevertheless, there are several approaches to uncertainty assessment that could potentially be employed for a variety of different modeling scenarios, such as the EPA guidelines to preparing economic analyses [27], the “ten commandments for good policy analysis” [26], and a survey of sampling-based methods for uncertainty and sensitivity analysis that includes a five step approach to uncertainty assessment [6].

Each of the above approaches touches on several of the requirements for a proper uncertainty assessment outlined in Section 1.3 and provides a good starting point for directing an uncertainty assessment. In general, these approaches consist of broad guidelines for conducting uncertainty assessment and tend to highlight a variety of methods that can be used for uncertainty classification and characterization, as well as different techniques for performing uncertainty and sensitivity analyses.

## 1.6 Uncertainty Assessment for Complex Models

As noted previously, current practices in uncertainty assessment are generally not specific to any particular type of numerical simulation model. As a result, most general approaches to uncertainty assessment suggest a variety of different methods for conducting uncertainty and sensitivity analyses rather than make specific recommendations [6, 26]. As will be discussed in Section 1.7, the objective in this work is not to present a variety of different methods for uncertainty assessment for general numerical simulation models, but rather, it is to present a specific, step-by-step approach to uncertainty assessment for complex models intended to support decision- and policy-making processes.

For these types of models, it is typically necessary to split the uncertainty as-

assessment into uncertainty assessment for decision-making support and uncertainty assessment for furthering the development of the models [1]. These types of models also tend to be computationally intensive, making sampling-based approaches to uncertainty assessment infeasible at times. For these situations, the uncertainty assessment of complex models requires the development of surrogate models on which the uncertainty assessment may be performed. Knowledge of the quantitative impacts on the analyses associated with the use of surrogate models in place of full models is essential to producing defensible claims in the context of decision-making in these cases. A more complete discussion of the needs for an uncertainty assessment for complex models is given in Chapter 2. What follows is a precise statement of the research objectives of this work.

## 1.7 Research Objectives

As mentioned earlier, most approaches to uncertainty assessment for complex models provide general guidelines rather than recommend specific methods and procedures. Thus, the goal of this research is to establish a specific approach to uncertainty assessment in a manner that is both broadly applicable and can be demonstrated on a real-world system. More specifically, the objectives of this research are,

1. To *establish* a probabilistic approach for assessing uncertainty in complex models intended to support decision-making and policy-making processes.
2. To systematically *develop* surrogate models for situations where proper assessment of uncertainty is computationally prohibitive.
3. To *demonstrate* the application of the general approach and surrogate modeling methodologies on real-world models designed to support decision-making and policy-making processes.

## The FAA Environmental Tools-Suite

The real-world application considered is a suite of tools being developed for the U.S. Federal Aviation Administration (FAA) Office of Environment and Energy, as well as NASA and NavCanada, which is intended to estimate the impacts of aviation on the environment. The goal of the development of this suite of tools is to establish the capability to characterize and quantify interdependencies among aviation-related noise and emissions, impacts on health and welfare, and industry and consumer costs, under different policy, technology, operational, and market scenarios. A key priority of the effort is to inform the analyses conducted by this suite of tools with the associated uncertainty from the factors and assumptions used in the analysis process [1].

The tools-suite, shown in Figure 1-3, consists of four main modules that model new technology and economics, simulate aircraft operations, and estimate the impacts of aviation on the environment. New technology is modeled with the Environmental

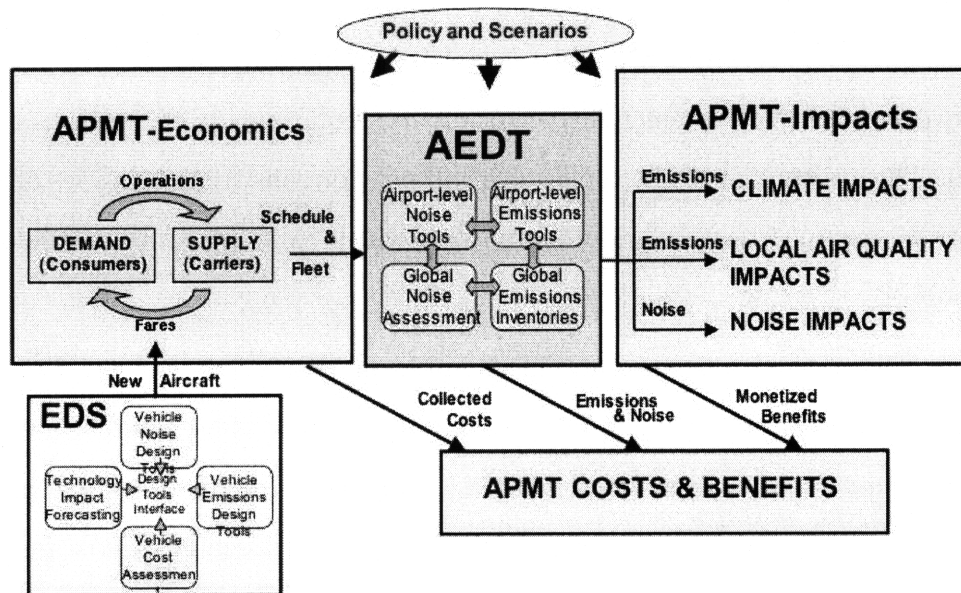


Figure 1-3: FAA tools-suite for estimating the impacts of aviation on the environment [1]

Design Space tool (EDS). EDS consists of a vehicle-level trade space where new technologies are identified and made available for fleet consideration. Economics are modeled with the Aviation environmental Portfolio Management Tool (APMT)

Economics block, which models the effects of policy scenarios on demand and accounts for the effects with growth, retirement, and replacement curves that model airline behavior. The simulation of aircraft operations is done by the Aviation Environmental Design Tool (AEDT), which consists of four models that map aircraft to operations, model operations to calculate thrust and fuelburn, estimate emissions production, and estimate noise produced. The impacts of aviation on the environment are accounted for in the APMT-Impacts block, which evaluates environmental impacts of aviation in terms of local air quality, climate change, and noise, with respect to physical, health-related, and monetary terms. The costs from the APMT-Economics block, noise and emissions from AEDT, and environmental impacts from the APMT-Impacts block can then be used for cost-effectiveness and cost-benefit calculations.

Each tool within the system contains a variety of uncertain factors and modeling assumptions that will impact both individual tool outputs, as well as overall system outputs. Thus, as mentioned previously, a critical component of the U.S. FAA tools-suite development program is a quantitative uncertainty analysis aimed at assessing the performance of the system relative to fidelity requirements for various analysis scenarios. The assessment of the tools-suite will provide sensitivity analyses of system outputs to factor uncertainties and assumptions, which will direct future development of the tools [1].

The models of the tools-suite are all complex, computationally intensive, and in some cases contain both aleatory and epistemic uncertainty. The different modules of the suite exhibit a wide-variety of disciplines, such as economics, aerodynamics, and atmospheric modeling, and substantially different computational needs and model characteristics, (e.g. AEDT is a database driven tool that can take days to run, while the climate module within the APMT-Impacts block is algorithm based and can take only a few hours to run). The development of an approach for uncertainty assessment and the systematic development of surrogate models for situations where analyses are computationally prohibitive, for this suite of tools, will thus demonstrate broad applicability of the methods and provide confidence that the approach can be extended to other complex models.

## 1.8 Thesis Outline

Chapter 2 presents the probabilistic approach established for assessing uncertainty in complex models. Chapter 3 focuses on the aspect of sensitivity analysis within the general approach. Chapter 4 discusses surrogate modeling methodologies and presents the creation of a surrogate model for uncertainty assessment of a real-world model. Chapter 5 presents the application of the full general approach on a real-world model. Chapter 6 contains general conclusions, as well as a discussion of future work.



# Chapter 2

## A Probabilistic Approach to Uncertainty Assessment

The first objective of this research is the establishment of a probabilistic approach to uncertainty assessment for complex models intended to support decision-making and policy-making processes. This chapter presents the approach developed in this work, as well as a detailed discussion of uncertainty characterization and uncertainty analysis.

### 2.1 Proposed Approach to Uncertainty Assessment

The proposed approach to uncertainty assessment follows the seven steps outlined in Section 1.3. Each step, as it relates to the uncertainty assessment of complex models, is discussed briefly in the following subsections, and Steps 4 and 5 are discussed in detail in Sections 2.2 and 2.3 respectively. Step 6 is discussed further in Chapter 3.

#### 2.1.1 Approach Step 1: Establish the desired outcomes of the uncertainty assessment

The first step given in the proposed seven-step approach is to establish a set of goals for the assessment. For complex models intended to support decision- and policy-making

processes, such goals should include both goals based on supporting decision-making and model development oriented goals. A proposed set of uncertainty assessment goals for Step 1, are as follows:

- Goals for supporting decision-making processes
  1. Provide quantitative evaluation of the performance of the model relative to fidelity requirements for various analysis scenarios.
  2. Provide quantitative comparisons of various policy scenarios, taking into account uncertainty in model outputs.
- Goals for furthering the development of the model
  1. Identify gaps in functionality that significantly impact the achievement of model requirements, leading to the identification of high-priority areas for further development.
  2. Rank factors based on contributions to output variability to inform future research and validation efforts.

As written, the goals are general, and are meant to sit above the entire uncertainty assessment process. Once the overarching set of outcomes has been established, whether they be for decision-making needs, model development, or both, a more detailed set of outcomes should be established as an internal set of questions that will lead to successfully meeting the goals of the assessment. A proposed set of questions, which when answered will meet the stated goals for supporting decision-making and for furthering the development of a model, is as follows:

- Questions for supporting decision-making processes:
  1. What are the key assumptions employed in the model?
  2. How do assumptions/limitations impact the applicability of the model for certain classes of problems?
  3. How do uncertainties in model factors propagate to uncertainties in model outputs?



4. For assumptions, limitations, and factors where effects cannot be quantified, what are the expected influences (qualitatively) on model outputs?
  5. How do uncertainty assessment results translate into guidelines for use?
- Questions for furthering model development:
    1. What are the key factors that contribute to variability in model outputs?
    2. Is there a need to direct research efforts at reducing output variability?
    3. If necessary, on which factors should research aimed at reducing output variability focus?

### **2.1.2 Approach Step 2: Documenting assumptions and limitations of the model**

For complex models intended to support decision- and policy-making, there will typically be many modeling assumptions employed, as well as inherent limitations to the model's capability. An example of an assumption in a model is an assumed discount rate for the estimation of a net present value for a particular policy scenario. An example of a limitation in a model is the inability of a model to analyze certain policies, such as the APMT-Impacts climate model's inability to study regional impacts of aviation [3].

For a model that is intended for use in support of decision- and policy-making processes, the transparent presentation of how each assumption impacts a model's performance, as well as limitations in terms of model applicability to certain classes of problems, is critical. For this task, assumptions and limitations should be documented with the following information:

- A description of the assumption, including what it means, where it enters the model, and what models/algorithms/databases it affects.
- The impacts of the assumption in terms of model validity (including references to previous validation work), and in cases where validation efforts have not taken place, a description of what it will take to assess the validity of an assumption.

- Implications for applicability of the module in terms of known policies and other uses for the model.

This information is pertinent to several of the development and decision-making oriented questions given in Section 2.1.1.

### **2.1.3 Approach Step 3: Documenting factors and outputs of the model**

Given that many factors of a complex model will have some degree of variability associated with them, it is necessary to establish what is known regarding the uncertainty introduced by each factor prior to determining how the uncertainty should be represented. The outputs of the model should also be identified at this point, as well as the factors on which they depend. This information is necessary for determining the type of uncertainty associated with each output, which is necessary for properly presenting the results of any analysis. For this task, factors and outputs should be documented with the following information:

- Factors
  - A description of the factor, including its units and what outputs it affects.
  - A description of available uncertainty information and the source of that information.
- Outputs
  - A description of the output, including its units, and possible downstream use.
  - A list of factors on which each output depends.

This information is critical for Step 4 of the approach, which is choosing how to classify and characterize uncertainties.

### 2.1.4 Approach Step 4: Classify and characterize uncertainty

As noted in Chapter 1, uncertainty is typically classed as aleatory or epistemic, where aleatory uncertainty arises through natural randomness and epistemic uncertainty arises through imperfect knowledge. Some studies decompose the epistemic uncertainty into epistemic uncertainty due to modeling choices and epistemic uncertainty used in the characterization of quantities known to contain aleatory uncertainty [33]. This decomposition of epistemic uncertainty leads to a hierarchical approach to uncertainty analysis, where epistemic modeling uncertainties are sampled in an outer loop, and aleatory and epistemic uncertainties relating to aleatory quantities are sampled in an inner loop. However, models, as they have been defined in this work, do not contain epistemic modeling uncertainties. Recall Section 1.2, where a model is defined as: a specific set of parameters and operations used in the representation of a system developed for the purpose of studying that system. The specific set of parameters alluded to in the definition can be considered a specific realization of epistemic modeling uncertainties, such as future scenarios. Within the FAA tools-suite discussed in Section 1.7, combinations of interest of epistemic modeling uncertainties constitute *lenses*, through which various policy scenarios are analyzed. Thus, the uncertainty assessment approach established in this work deals with only aleatory and epistemic uncertainty that relate to model factors, rather than on modeling assumptions or uncertainties. The notion of a lens and the relation between lenses and epistemic modeling uncertainty are discussed in more detail in Chapter 5.

For complex models intended to support decision-making, the proposed method for characterizing factor uncertainty is:

- aleatory factors: use natural probabilistic representation,
- epistemic factors: use maximum uncertainty principle [20] to define a probabilistic distribution based on uncertainty information,

A more detailed account of how uncertainty should be characterized for complex models intended to support decision- and policy-making processes is delayed until Section 2.2.

### **2.1.5 Approach Step 5: Conduct uncertainty analysis**

Uncertainty analysis encompasses the process of characterizing and analyzing the effects of uncertainty in model factors, with a focus on quantitative assessment of the effects on model outputs and thus, on the conclusions drawn from simulation results. This step of the approach, which will be elaborated on further in Section 2.3, uses the information gained in the previous four steps to answer the questions given in Section 2.1.1 for supporting decision-making.

### **2.1.6 Approach Step 6: Conduct sensitivity analysis**

Sensitivity analysis studies how variability in model outputs can be apportioned to sources of uncertainty in model factors [5]. For complex models intended to support decision- and policy-making processes, this step in the uncertainty assessment is critical for situations where output variability, as determined through uncertainty analysis, is too substantial for model results to be useful. These situations lead to the need to consider the epistemic uncertainties present to determine where research should be focused to help trim output variability. Properly classifying uncertainties is thus critical to sensitivity analysis for complex models used to support decision-making. The importance of making the distinction between uncertainty types is made clear by the following illustrative example.

Consider two fair coins, labeled coin A and coin B, that are tossed independently at the same time and immediately covered so that it is unknown what the results of the tosses were. Assume that a coin that lands on heads is a success and a coin that lands on tails is a failure. A model that then estimates the number of successes in this experiment would estimate the expected number of successes as 1, and the variance of the number of successes as 0.5.

Now consider a second experiment, which is the same as the first experiment with the following modification: after the coins are tossed, an onlooker notices the result of the coin A toss before the coin can be covered. Assuming this onlooker cannot be consulted, the expected number of successes and the variance of the number of

successes would be evaluated as they were for the first experiment. However, if the onlooker can be consulted, then the result of the coin A toss will be known. Assume coin A landed on heads, and was thus a success. Using the same modeling method as in the first experiment, the number of successes would then be estimated as 1.5, and the variance of the number of successes would be estimated as 0.25. The difference in this case, is that the outcomes of the first experiment, with equal probability, were  $\{HH, HT, TH, TT\}$  or  $\{2, 1, 1, 0\}$ , where the coins are ordered A,B, whereas in the second experiment, the outcomes, again with equal probability, were  $\{HH, HT\}$  or  $\{2, 1\}$ .

This example could be taken one step further, to the point where the onlooker sees the result of both tosses, in which case the number of successes would be known precisely, and the variance of the estimate of the number of successes would become zero. This simple example illustrates the notion of reducible uncertainty, and thus the difference between aleatory and epistemic uncertainty. In the first experiment, the results of the coin tosses could not be known, and once the coins were covered, could never be known. In this case the uncertainty associated with the outcome of the coin tosses was associated entirely with the random process of tossing a fair coin. In the second experiment, information regarding the result of one of the tosses was known. In this experiment, the uncertainty associated with coin A was epistemic. Without consulting the onlooker, imperfect knowledge was had of the result of the toss of coin A, while coin B still contained aleatory uncertainty. Maintaining the imperfect knowledge led to the same conclusions in the second experiment as were found in the first, however, the variability of the second experiment could be reduced by attempting to gain better information about coin A. Once the knowledge of coin A is obtained, the variability of the estimate of the number of successes in the second experiment is reduced to half of that in the first experiment.

This notion of reducible uncertainty is prevalent in complex models intended to support decision- and policy-making processes. Thus, the identification of which uncertainties are reducible, which is not always as straightforward as it is for the coin tossing example and may be related to the current state of knowledge, and the

amount by which they can be reduced are important aspects in sensitivity analysis. Sensitivity analysis aimed at meeting the goals given in Step 1 of the approach is discussed in detail in Chapter 3.

### **2.1.7 Approach Step 7: Present results**

The final step of the proposed approach is the preparation and presentation of results. As noted in Section 1.4, the visual presentation of quantitative information, such as the results of uncertainty and sensitivity analyses, is a distinct area of research, and it is recommended that results presentations adhere to the recommendations found in [31, 32].

Though the approach to uncertainty assessment for complex models intended to support decision- and policy-making processes concludes with this step, it is important to note that the information provided by the assessment, particularly uncertainty analysis, is intended to be used in support of decision-making, though the results are not in the form of an evaluated decision-rule or utility function. Given that there is a vast body of literature related to formal theories of decision-making that include the identification of values, risk-aversion, hidden objectives, alternatives, and many other attributes that are involved in the formation of a utility function that may be used to compare a variety of different policy options [34, 35, 36], it is an objective of this work only to support this process, not direct it, with uncertainty assessment results.

## **2.2 Characterizing Uncertainty**

In general, there is a disconnect between theoretical characterization and the applied treatment of uncertainty in sensitivity and uncertainty analyses. Theoretical characterizations of uncertainty, such as classifying uncertainty as being aleatory or epistemic, are typically not treated differently in applications of uncertainty assessment. For example, most practitioners of global sensitivity analysis, the state-of-the-art method of variance apportionment, which will be discussed in detail in Section 3.2, treat all uncertainty as if it is epistemic [30]. Another example is the widespread

use of Monte Carlo simulation for propagating factor uncertainty to model outputs, where factor uncertainty is treated implicitly as aleatory, thus causing outputs to inherit the supposed aleatory nature of the uncertainty.

As discussed in Section 1.4, to maintain theoretical uncertainty characterization in the applied treatment of uncertainty using sampling-based probabilistic methods, it is necessary to characterize uncertainty probabilistically, in a manner that is consistent, meaningful, and defensible. In the context of decision- and policy-making applications, models are typically used in ampliative reasoning, that is, problems involving drawing conclusions that are not entailed in the given premises. For this type of reasoning, it is essential that uncertainty be characterized via the *principle of maximum uncertainty* [20]. In the case of complex models, the premises are the uncertainty information associated with model factors, such as ranges and most-likely values, and the conclusions are the uncertainty associated with model outputs and any decisions made using that information. When employing the probabilistic method for representing uncertainty, assigning a probability distribution to a given factor is in fact implying that more is known about the uncertainty associated with that factor than is known from the information at hand. The propagation of this uncertainty through a model to model outputs can then lead to estimates of output probability distributions, which gives the appearance of fully quantified uncertainty. The purpose of employing the principle of maximum uncertainty is to avoid drawing conclusions based on information not contained in given premises by maximizing nonreliance on information not contained in premises [20].

The principle of maximum uncertainty is enforced by selecting probability distributions that maximize some measure of uncertainty. Typical uncertainty measures are quantities such as the standard deviation of a random variable and the entropy of a random variable [20, 37]. For establishing maximum uncertainty distributions, maximizing entropy produces reasonable results, whereas maximizing standard deviation does not (i.e. maximizing the standard deviation of a random variable with the information that the factor takes on values in a given range would place all of the probability density at the minimum or maximum points of the interval). Thus,

Constraints	Maximum Entropy Distribution
Minimum value = a Maximum value = b	Uniform
Expected value = $\mu$ Standard deviation = $\sigma$	Normal
Expected value = $\mu$ Standard deviation = $\sigma$ Minimum = a Maximum = b	Beta
Discrete values	Discrete uniform

the measure of uncertainty recommended for assigning maximum uncertainty distributions is information theoretic entropy. Entropy is defined as

$$H(X) = - \sum_{i=1}^n p(x_i) \log p(x_i), \quad (2.1)$$

for the case of discrete random variables, where  $H(X)$  is the entropy,  $X$  is some discrete random variable,  $p(x_i)$  is the probability that  $X = x_i$ , and there are  $n$  possible values  $x$  can take [37]. For the continuous case, entropy can be defined as

$$h(X) = - \int_{\mathbb{X}} f(x) \log f(x) dx, \quad (2.2)$$

where  $h(X)$  is the entropy,  $X$  is some continuous random variable,  $\mathbb{X}$  is the support of  $X$ , and  $f(x)$  is the probability density function of  $X$  [37].

Maximum entropy distributions, and thus distributions satisfying the principle of maximum uncertainty, are given in Table 2.1 for common constraints regarding uncertainty information [20]. The distributions assigned by the principle of maximum uncertainty for common situations of uncertainty information given by the constraints are appealingly, well-known distributions, and in most applications of probabilistic



uncertainty methods, are used frequently. Each of these distributions is discussed in more detail below.

A uniform distribution can be defined by two real numbers,  $a$ ,  $b$ , such that  $a < b$ . A random variable, say  $X$ , with such a distribution is said to be distributed uniformly on the interval  $[a, b]$ , or  $X \sim U[a, b]$ . The probability density function for a uniform random variable is then given as:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{for } a \leq x \leq b, \\ 0 & \text{otherwise.} \end{cases}$$

Figure 2-1 shows an example of a uniform distribution on the interval  $[0, 1]$ . A factor

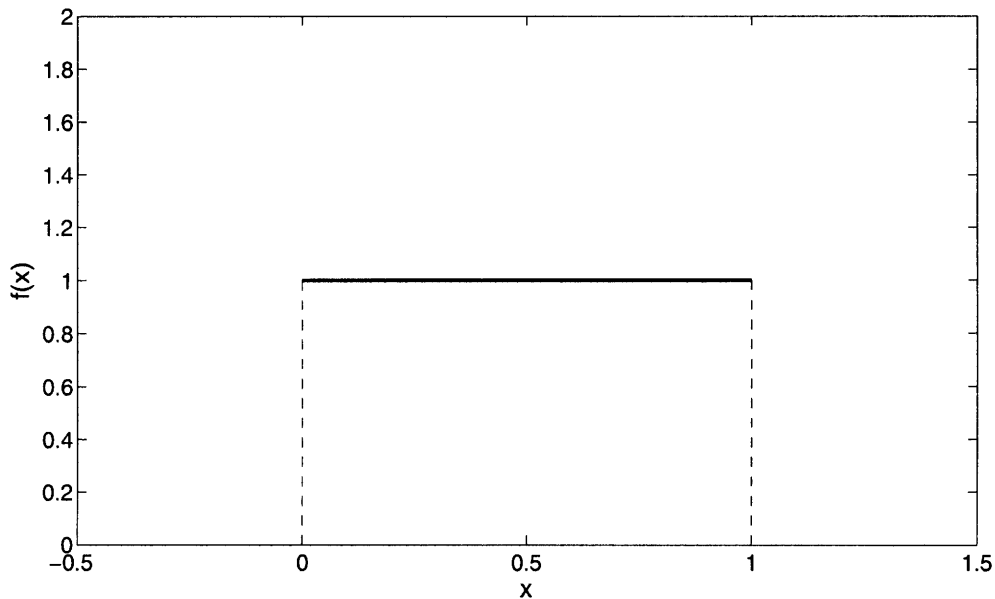


Figure 2-1: Example of a uniform probability density function.

that has been assigned a uniform distribution, which in the case of the principle of maximum uncertainty is a factor that is only known to take on values on some interval, is assigned an equal probability of taking on a value in any equally sized interval on  $[a, b]$ .

A discrete uniform distribution is a probability distribution for which all values of a finite set of possible values have equal probability. For example, if a factor can

take the values,  $k_1, \dots, k_m$ , then each will have a probability of  $1/m$  of occurring. The probability mass function is thus,

$$f(x) = \frac{1}{m} \text{ for } x \in \{k_1, \dots, k_m\}. \quad (2.3)$$

Figure 2-2 shows an example of a discrete uniform distribution for the case where  $x$  can take the values in the set  $\{k_1, \dots, k_{10}\}$ .

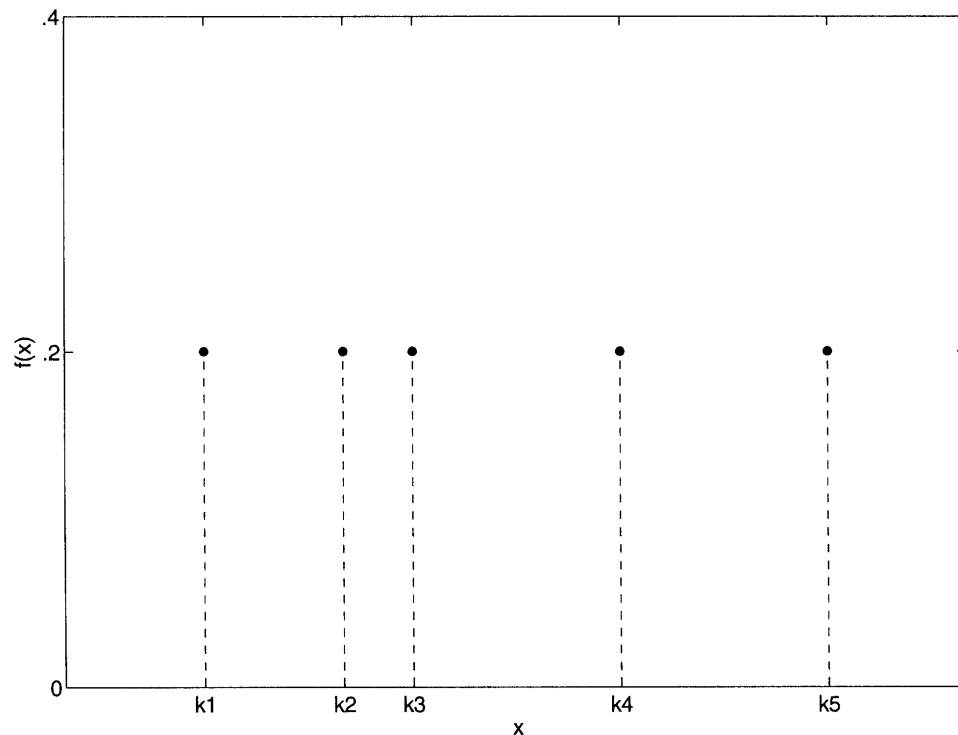


Figure 2-2: Example of a discrete uniform probability mass function.

A normal distribution, which is also referred to as a Gaussian distribution, is typically defined by a mean,  $\mu$ , and a variance,  $\sigma^2$ . A random variable,  $X$ , with such a distribution is said to be distributed normally with mean,  $\mu$ , and variance,  $\sigma^2$ , or  $X \sim \mathcal{N}(\mu, \sigma^2)$ . The probability density function for a normal random variable is

given as:

$$f(x|\mu, \sigma^2) = \frac{1}{(2\pi)^{1/2}\sigma} \exp \left[ -\frac{1}{2} \left( \frac{x - \mu}{\sigma} \right)^2 \right] \text{ for } -\infty < x < \infty. \quad (2.4)$$

Figure 2-3 shows an example of a normal distribution with  $\mu = 0$  and  $\sigma^2 = 1$ .

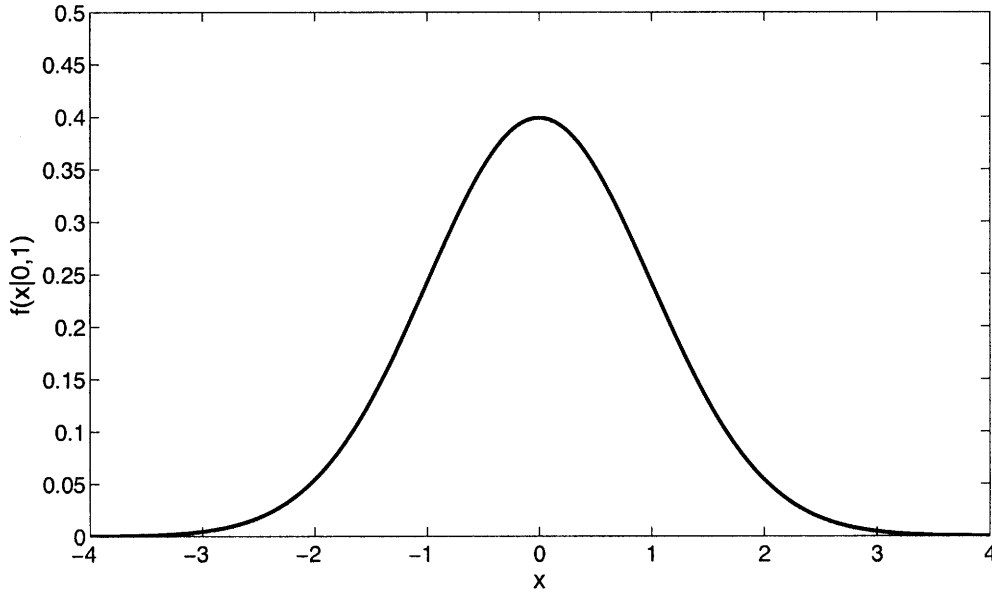


Figure 2-3: Example of a normal probability density function.

A beta distribution is typically defined for the interval  $[0, 1]$  with two parameters,  $\alpha$  and  $\beta$ , that define the shape of the distribution. The probability density function for a beta distribution is given as:

$$f(x|\alpha, \beta) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1} & \text{for } 0 \leq x \leq 1, \\ 0 & \text{otherwise,} \end{cases}$$

where  $\Gamma(\cdot)$  is the gamma function, which is defined as

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx. \quad (2.5)$$

An example of the beta distribution for  $\alpha = 2$ ,  $\beta = 2$  is shown in Figure 2-4. The

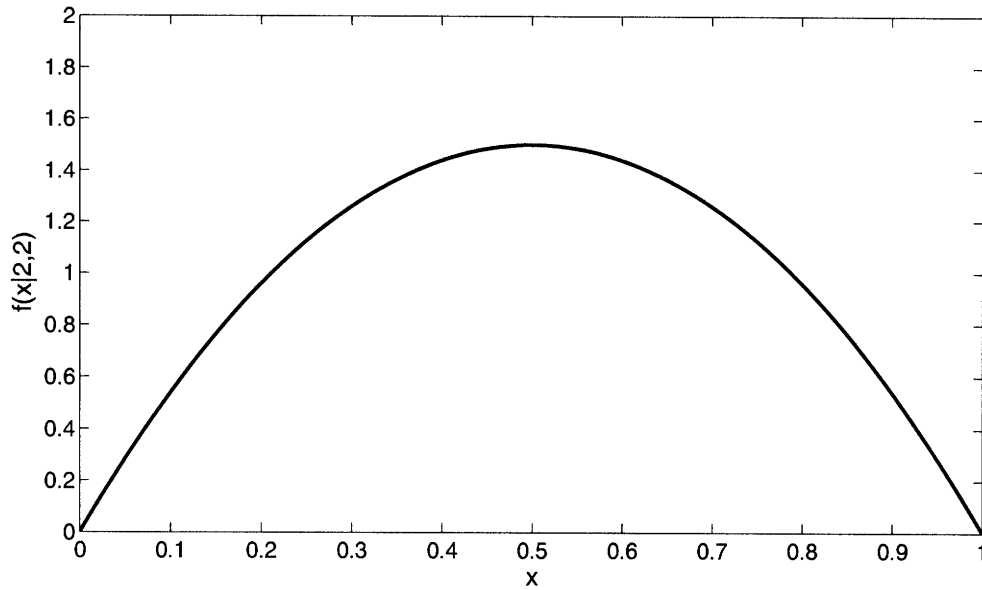


Figure 2-4: Example of a beta probability density function.

beta distribution is a useful distribution in the sense that its shape parameters give it a lot of flexibility, and it is defined on a finite interval rather than on  $(-\infty, \infty)$ . However, as can be seen from the definition of its probability density function, a beta distribution is complex in that the parameters of the distribution do not have obvious interpretations. For this reason, it is common in modeling for decision- and policy-making to use a triangular distribution as a proxy for a beta distribution [38, 39]. The triangular distribution is a more understandable means for quantifying uncertainty than a beta distribution in the sense that the role of the parameters in this family of distributions is transparent. A triangular distribution is typically defined by three parameters, a minimum,  $a$ , a maximum,  $b$ , and a mode,  $c$ . A random variable,  $X$ , with such a distribution is said to be triangularly distributed with parameters  $a, b$  and  $c$ , or  $X \sim T(a, b, c)$ . The probability density function for a triangular random variable is given as:

$$f(x|a, b, c) = \begin{cases} \frac{2(x-a)}{(b-a)(c-a)} & \text{for } a \leq x \leq c, \\ \frac{2(b-x)}{(b-a)(b-c)} & \text{for } c \leq x \leq b, \\ 0 & \text{otherwise.} \end{cases}$$

An example of a triangular distribution with parameters,  $(-1, 1, 0)$ , is shown in Figure 2-5.

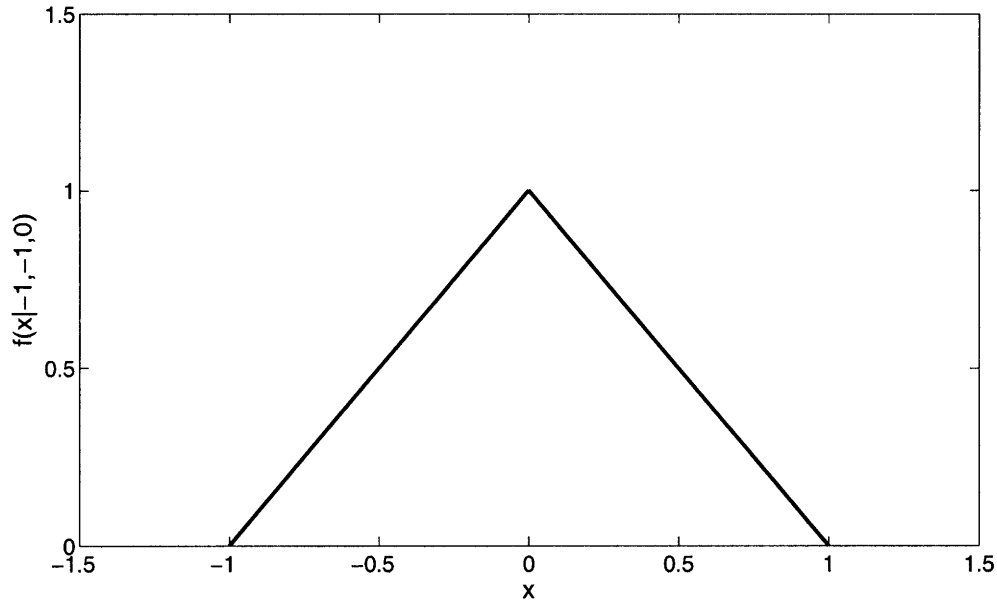


Figure 2-5: Example of a triangular density function.

Thus, for factors with epistemic uncertainty, distributions defined by the maximum uncertainty principle should be used to characterize uncertainty, however, the triangular distribution is recommended in place of the beta distribution. As noted in Section 2.1.4, factors with aleatory uncertainty should use their natural probabilistic representations.

In this work, sampling-based probabilistic approaches are considered, however, it should be noted, as it was in Section 1.3, that there are a variety of different methods for representing and characterizing uncertainty, such as possibility theory, evidence theory, interval analysis, and many others, and it is important to understand that the results and inferences that can be made from a particular study depends on the treatment of the uncertainty in the analyses [40]. The representation and characterization of uncertainty in this work follows a Bayesian perspective in its application of the principle of maximum uncertainty, where probability is used as a logic for reasoning rather than as a means of assigning frequencies to certain events [41]. The benefit of

this approach is that it permits the use of the mathematical methods of probability theory in uncertainty assessment, however, it is important that results not be interpreted from a frequency perspective. In Section 5.2.5, the implications of using the principle of maximum uncertainty to represent epistemic uncertainty are discussed further, as well as considerations of how a non-probabilistic method, such as interval analysis, could have been employed and interpreted.

## 2.3 Uncertainty Analysis

As noted in Section 2.1.5, the purpose of uncertainty analysis in the approach is to answer the questions for supporting decision-making. The key question answered by uncertainty analysis is, “how do uncertainties in model factors propagate to uncertainties in model outputs?” The answer to this question, which is typically given in terms of model output means, variances, and the construction of output histograms, can be used to provide quantitative comparisons of various policy scenarios and quantitative evaluation of the performance of the model relative to fidelity requirements.

The computation of model output means, output variances, and other distributional information in support of uncertainty analysis for decision-making can be carried out with several different methods, such as mean-value methods [42], analytic reliability methods [43], stochastic expansion methods, such as polynomial chaos [44], and sampling-based techniques referred to as Monte Carlo simulation [21, 22].

Mean-value methods consider estimations (usually first or second-order) of model outputs around mean-value points. These methods are computationally inexpensive but can be problematic in the presence of non-linear model responses. Analytic reliability methods, such as the first- and second-order reliability methods (FORM and SORM), use analytic methods to approximate the probabilities of certain events (usually some form of failure event). The methods are computationally efficient relative to sampling-based approaches for linear models, however, this efficiency is lost in the case of non-linear black-box models [45]. Stochastic expansion methods, such as polynomial chaos expansions, model the relationship between outputs and factors

using orthogonal polynomials for which the uncertainty analysis process is reduced to the estimation of the coefficients of the orthogonal polynomials. Methods such as polynomial chaos expansions are used generally in situations where model structure is known, such as systems governed by partial differential equations.

Though a variety of methods exist for uncertainty analysis, such as those discussed above, as noted in Section 1.3, only sampling-based approaches are considered here due to their general applicability, effectiveness, and wide use. These methods are discussed in the following paragraphs.

Consider a general model  $f(\mathbf{x})$ , where  $\mathbf{x} = [X_1, X_2, \dots, X_k]^T$  is the vector of  $k$  factors of the model. If the model factors are viewed as random variables with some associated probability distribution, then the mean value of the model output can be computed from a Monte Carlo simulation as

$$\frac{1}{N} \sum_{m=1}^N f(\mathbf{x}^m) \rightarrow \mathbb{E}[f(\mathbf{x})] \text{ as } N \rightarrow \infty, \quad (2.6)$$

where  $N$  is the number of model evaluations in the Monte Carlo simulation and  $\mathbf{x}^m = [X_1^m, X_2^m, \dots, X_k^m]^T$  denotes the  $m$ th sample realization of the random vector  $\mathbf{x}$ . Convergence of the sample mean in Equation 2.6 to the expected value of  $f(\mathbf{x})$  is guaranteed by the law of large numbers and the convergence rate is  $1/\sqrt{N}$ , as given by the Central Limit Theorem [46]. Output variances and other distributional quantities can similarly be computed using Monte Carlo simulation results.

There are several different sampling strategies that can be used to evaluate expressions such as the one shown in Equation 2.6. The most common methods are brute force pseudorandom sampling, quasi-Monte Carlo sampling [47], and Latin hypercube sampling [48]. Brute force pseudorandom sampling consists of selecting samples of factors randomly from their probability distributions. The method is referred to as pseudorandom sampling because a computer's pseudorandom number generator is typically used to generate the samples. When using a pseudorandom number generator it is important to be sure that it has been tested and verified using for example, the diehard battery of tests of randomness [49]. Quasi-Monte Carlo sampling selects

samples of factors deterministically using what are referred to as low-discrepancy sequences that aim to sample a space as uniformly as possible. In high-dimensions, these methods tend to have problems with some factors being highly correlated with other factors, and thus, care should be taken in the application of quasi-Monte Carlo sampling for high-dimensional models. Latin hypercube sampling is a method of selecting samples of factors in a manner that ensures all factors have been sampled across their entire domains. The advantages of Latin hypercube sampling are greatest when the number of samples is small ( $\mathcal{O}(100)$ ), and diminishes as the number of samples increases. Each method can dominate the other methods in terms of the number of samples required to achieve equally accurate estimates under certain circumstances, thus, the best sampling strategy depends on the model and the quantity being estimated. For the uncertainty assessment approach presented here, brute force pseudorandom sampling is recommended for uncertainty analysis since it is the most general method and it is anticipated that many complex models will be high-dimensional and require many samples. Further, brute force pseudorandom sampling is currently required for aspects of sensitivity analysis that are discussed in Chapter 3.



# Chapter 3

## Sensitivity Analysis for Complex Models

The purpose of conducting a sensitivity analysis in the approach to uncertainty assessment established in this work is to meet the goals of the assessment for model development purposes by answering two of the model development questions given in Section 2.1.1. Those questions are:

- What are the key factors that contribute to variability in model outputs?
- On which factors should research aimed at reducing output variability focus?

Knowledge of the key factors that contribute to variability in model outputs serves the purpose of a “sanity check” in terms of model validity. If certain anticipated key factors are not identified as significant contributors, then future development efforts can focus on further model verification and validation exercises. If the identified key factors are as anticipated, further confidence in the validity of the model is gained. For situations where output variability is so large that model results are useless for supporting decision-making, knowledge of which factors should be researched further to reduce output variability is essential to the future application of the model.

This chapter presents background material on available methods for conducting sensitivity analysis for complex models in Section 3.1, followed by a detailed discussion of global sensitivity analysis in Section 3.2. Following that discussion is the

development of an original method, referred to as distributional sensitivity analysis, in Section 3.3, which I developed to answer the model development question regarding focusing future research aimed at reducing output variability.

### 3.1 Sensitivity Analysis Background

There are several methods that can be used for answering the questions aimed at meeting the goals of uncertainty assessment for model development purposes. Among the most common methods are iterated fractional factorial design (IFFD) [50], the standardized regression coefficients (SRC) [51], the Spearman rank correlation test [52], vary-all-but-one analysis (VABO) [3], and global sensitivity analysis [5]. The IFFD method is based on design of experiments techniques. The method establishes key drivers of output variability by sampling factors with a resolution IV orthogonal fractional factorial design aimed at determining both linear and quadratic impacts of each factor on model outputs. However, the method is not capable of exploring any higher-order effects, and is thus not a rigorous quantitative means of apportioning output variance [50]. The SRC method proceeds by conducting a Monte Carlo simulation and using the results of the simulation to generate a least-squares regression model. The coefficients for each factor in the regression model are standardized by the ratio of the standard deviation of the factor distribution and the standard deviation of the output. The SRC method explores only linear effects and is thus also not a rigorous means for answering the questions aimed at meeting the goals of uncertainty assessment for model development purposes. The Spearman rank correlation test proceeds by running a Monte Carlo simulation followed by a ranking of the factor and output samples based on their positions in their respective order statistics [52], meaning the smallest value in the samples of each factor and of the output is given the rank 1, the next smallest of each is given rank 2, etc. The Pearson product-moment correlation coefficient is then computed on the rank data for each factor with the output. Key drivers are then identified as those factors with large coefficients (in magnitude), which imply strong linear relationships with the output. Like the SRC method, the

Spearman rank correlation test is a linear method, and is thus not a rigorous means for apportioning output variance. The VABO method, unlike the previous methods discussed, is not linear in nature. The method proceeds by running a Monte Carlo simulation and computing output variance. Then, a particular factor is fixed to a point on its domain, and another Monte Carlo simulation is conducted. The difference between the variance of the first Monte Carlo simulation and the second, is considered the contribution of the fixed factor to output variability. This process is repeated for each factor in the model. The key drawback to this method of apportioning output variance is that it is not obvious where each factor should be fixed on its domain, which can lead to a variety of different variance apportionments depending on how the factors are fixed, and can even at times lead to situations where fixing a given factor increases output variability [5]. Thus, VABO methods are also not a rigorous means for answering the questions aimed at meeting the goals of uncertainty assessment for development purposes. The method of global sensitivity analysis is an extension of the VABO method that takes into account all possible locations each factor can be fixed on their domains. As a result, it is considered a rigorous method for quantitatively apportioning output variance [17], and is recommended here for identifying the key factors that contribute to output variability. This method is discussed in detail in the following section. An original method that makes use of global sensitivity analysis results for answering the third question for development uncertainty assessment given in Section 2.1.1 is discussed in detail in Section 3.3.

## 3.2 Global Sensitivity Analysis

The goal of a global sensitivity analysis is shown notionally in Figure 3-1, where the pie represents the variance in a model output, which is then decomposed according to factor contributions. The results of a global sensitivity analysis permit a ranking of model factors that can be used in different development settings such as *factor prioritization* for future research, where the goal is to determine which factors, once fixed will cause the largest reduction in variance, and *factor fixing*, for which the goal

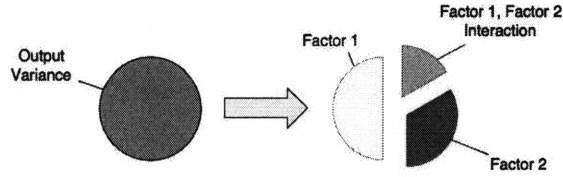


Figure 3-1: Apportioning Output Variance

is to identify noninfluential factors that may be fixed without substantially affecting model outputs [5].

The process of apportioning output variance across model factors in a global sensitivity analysis can be carried out by both a Fourier Amplitude Sensitivity Test (FAST) method, and the Sobol' method [17, 18, 19, 53]. The FAST method is based on Fourier transforms, while the Sobol' method utilizes Monte Carlo simulation. The Sobol' method is discussed here.

The Sobol' method for computing global sensitivity indices was proposed by Russian mathematician I.M. Sobol'. The method is well-developed and in wide use in the sensitivity analysis field, particularly by the Joint Research Centre of the European Commission [5, 17]. The method is discussed here in detail because the surrogate modeling methods developed in Chapter 4 will make use of the formulation. The derivation follows the work of Homma and Saltelli [18].

The Sobol' method is based on the ANOVA High-Dimensional Model Representation (ANOVA-HDMR). A high-dimensional model representation of a function,  $f(\mathbf{x})$ , can be written as

$$f(\mathbf{x}) = f_0 + \sum_i f_i(x_i) + \sum_{i < j} f_{ij}(x_i, x_j) + \dots + f_{12\dots n}(x_1, x_2, \dots, x_n), \quad (3.1)$$

where  $f_0$  is a constant,  $f_i(x_i)$  is a function of only  $x_i$ ,  $f_{ij}(x_i, x_j)$  is a function of only  $x_i$  and  $x_j$ , etc. Without any constraints, the representation of  $f(\mathbf{x})$  given by Equation 3.1 is not unique, however, it can be made unique by enforcing the constraints

$$\int_0^1 f_{i_1, \dots, i_s}(x_{i_1}, \dots, x_{i_s}) dx_k = 0, \text{ for } k = i_1, \dots, i_s, \quad s = 1, \dots, n, \quad (3.2)$$

where the function  $f(\mathbf{x})$ , and hence all its components, has been assumed to be integrable. For simplicity of presentation, the factors of the function in Equation 3.2 have been defined on the interval  $[0,1]$ , but this assumption is not essential to the method. For each  $s$ , the indices  $i_1, \dots, i_s$  in Equation 3.2 are all sets of  $s$  integers such that  $1 \leq i_1 < \dots < i_s \leq n$ . Thus, for  $s = 1$ , the constraint given by Equation 3.2 applies to all terms  $f_i$  in Equation 3.1, while for  $s = 2$ , the constraint Equation 3.2 applies to all terms  $f_{ij}$  with  $i < j$  as in Equation 3.1, etc. The application of the constraint Equation 3.2 makes the HDMR a unique representation of the function  $f(\mathbf{x})$ , referred to as an ANOVA-HDMR. Integration of  $f(\mathbf{x})$  over all inputs results in  $\int f(\mathbf{x})d\mathbf{x} = f_0$ , which assuming each input  $x_i$  is a uniform random variable on  $[0,1]$ , is the mean value of the function  $f(\mathbf{x})$ .

The constraint given by Equation 3.2 also forces the different components of  $f(\mathbf{x})$  within the ANOVA-HDMR to be orthogonal. That is, if  $(i_1, \dots, i_s) \neq (j_1, \dots, j_l)$ , then

$$\int f_{i_1, \dots, i_s}(x_{i_1}, \dots, x_{i_s}) \times f_{j_1, \dots, j_l}(x_{j_1}, \dots, x_{j_l}) d\mathbf{x} = 0, \quad (3.3)$$

since at least one index is not repeated.

Assuming now that  $f(\mathbf{x})$  is square integrable, and therefore all components within the ANOVA-HDMR are as well, the variance of  $f(\mathbf{x})$  is written as

$$D = \int f(\mathbf{x})^2 dx - f_0^2, \quad (3.4)$$

and partial variances are defined as

$$D_{i_1 \dots i_s} = \int f_{i_1, \dots, i_s}(x_{i_1}, \dots, x_{i_s})^2 dx_{i_1} \dots dx_{i_s}. \quad (3.5)$$

Given the ANOVA-HDMR for some  $f(\mathbf{x})$ , we square and then integrate both sides of Equation 3.1, and employ the orthogonality constraint to arrive at

$$\int f(\mathbf{x})^2 dx = f_0^2 + \sum_i D_i + \sum_{i < j} D_{ij} + \dots + D_{12 \dots n}, \quad (3.6)$$

which implies

$$D = \sum_i D_i + \sum_{i<j} D_{ij} + \dots + D_{12\dots n}. \quad (3.7)$$

This is precisely the notion shown in Figure 3-1.

Global sensitivity indices are defined as

$$S_{i_1, \dots, i_s} = \frac{D_{i_1 \dots i_s}}{D}, \quad s = 1, \dots, n. \quad (3.8)$$

The sum of all global sensitivities of this form for a given function is unity. Global sensitivity indices with only one subscript, (e.g.  $S_i$ ), are called main effect sensitivities, and those with multiple subscripts, (e.g.  $S_{i,j}$ ,  $S_{i,j,k}$ , etc.), are called interaction effect sensitivities. The sum of a factor's main effect global sensitivity and all interaction effect sensitivities that involve that factor gives the total effect sensitivity index,  $\tau$ , which is defined for input factor  $i$  as

$$\tau_i = S_i + S_{i,ic}, \quad (3.9)$$

where  $S_i$  is the main effect sensitivity to factor  $i$ , and  $S_{i,ic}$  is the sum of the sensitivity indices of all interaction effects that include factor  $i$ . Since the sum of all unique sensitivity indices is unity, we have that

$$\tau_i = S_i + S_{i,ic} = 1 - S_{ic}, \quad (3.10)$$

where  $S_{ic}$  is the sum of the sensitivity indices for all main effects and interactions effects that do not involve factor  $i$ . Since interaction effects will be counted for each factor involved in them,  $\sum_i \tau_i \geq 1$ .

The main and total effect sensitivity indices can be computed via Monte Carlo simulation as follows [18], where hat quantities denote estimates of the corresponding true quantities. Here it should be noted that the computation of the partial variances with Monte Carlo simulation proceeds directly with the function  $f(\mathbf{x})$  and does not require explicit knowledge of the functions on the right-hand side of Equation 3.1.

The estimate of the mean  $f_0$  is computed as

$$\hat{f}_0 = \frac{1}{N} \sum_{m=1}^N f(\mathbf{x}^m), \quad (3.11)$$

while the estimate of the variance  $D$  is

$$\hat{D} = \frac{1}{N} \sum_{m=1}^N f(\mathbf{x}^m)^2 - \hat{f}_0^2. \quad (3.12)$$

The single-factor partial variance is then computed for factor  $i$  by resampling all factors except factor  $i$ :

$$\hat{D}_i = \frac{1}{N} \sum_{m=1}^N f([x_1^m, \dots, x_i^m, \dots, x_n^m]^T) f([\tilde{x}_1^m, \dots, x_i^m, \dots, \tilde{x}_n^m]^T) - \hat{f}_0^2, \quad i = 1, \dots, n, \quad (3.13)$$

where  $\tilde{x}_j^m$  denotes a different sample of factor  $x_j$ . The main effect sensitivity index can then be estimated as,  $S_i = \hat{D}_i / \hat{D}$ . The estimate of the variance due to all factors except factor  $i$  (which includes the sum of all single-factor and interaction effect partial variances that do not include factor  $i$ ) is denoted as  $\hat{D}_{i^c}$ , and is computed by

$$\hat{D}_{i^c} = \frac{1}{N} \sum_{m=1}^N f([x_1^m, \dots, x_i^m, \dots, x_n^m]^T) f([x_1^m, \dots, \tilde{x}_i^m, \dots, x_n^m]^T) - \hat{f}_0^2, \quad (3.14)$$

where now just factor  $i$  is resampled. Finally, computing  $\hat{S}_{i^c} = \hat{D}_{i^c} / \hat{D}$  and applying Equation 3.10, we obtain the desired total effect sensitivity index.

The main effect sensitivity indices,  $S_i$ , may be used for factor prioritization by ranking factors according to their main effect indices, which give the percentage of how much output variability can be expected to be eliminated by fixing a particular input somewhere on its domain. The total effect sensitivity indices,  $\tau_i$ , may be used for factor fixing, since a low total effect index reveals a given input has a small main effect and also does not take part in substantial interactions among other factors. For  $n$  factors, the calculation of sensitivity indices requires  $(n+2)$  Monte Carlo simulations (each with  $N$  model evaluations) if both the main effect and total effect indices are

desired. Convergence of the estimates of main and total effect indices is discussed in Section 3.4.

### 3.3 Distributional Sensitivity Analysis

Global sensitivity analysis is a rigorous method for apportioning output variance, however, it is not generally a rigorous method for determining how to focus future research efforts aimed at reducing output variability. The key drawback to using global sensitivity analysis results in the factor prioritization setting (that is, to direct future research), is the underlying assumption that a given factor can, through further research, be fixed to some point on its domain. For epistemic factors, this is an optimistic assumption, which, as will be shown below, can lead to inappropriate allocation of resources. Further, for factors containing both aleatory and epistemic uncertainty, the assumption cannot be met.

To account for the inherent limitations in using global sensitivity analysis results for directing future research, an original method, which will be henceforth be referred to as *Distributional Sensitivity Analysis* has been developed. Rather than look at *factor prioritization* by considering which factors, once fixed, cause the greatest reduction in output variance, the method focuses on determining which factors would on average cause the greatest reduction in output variance, given that the portion of a particular factor's variance that can be reduced is a random variable. A key aspect of the method is that the analysis is performed directly on the factor and output samples that were generated during a global sensitivity analysis, thus the the cost remains at  $N(n + 2)$  model evaluations. The derivation of the method is given in the following subsections, which includes a discussion on acceptance/rejection sampling, which is the technique that permits the reuse of the samples from global sensitivity analysis.



### 3.3.1 Derivation of distributional sensitivity analysis

Consider a generic model

$$Y = f(\mathbf{x}), \quad (3.15)$$

where  $\mathbf{x} = [X_1, \dots, X_m]^T$ , and  $X_1, \dots, X_m$ , are random variables and thus,  $Y$  is a random variable as well. By definition, the total variance of a random variable  $Y$ , can be decomposed for any random variable  $X_i$ , as

$$\text{var}(Y) = \mathbb{E}[\text{var}(Y|X_i)] + \text{var}(\mathbb{E}[Y|X_i]). \quad (3.16)$$

Both  $\mathbb{E}[\text{var}(Y|X_i)]$  and  $\text{var}(\mathbb{E}[Y|X_i])$  are greater than zero and thus, both are also less than the total variance of  $Y$ . Therefore, when the expected value of the variance of an output given a particular factor is high, the variance of the expected value of the output given that particular factor is low, and vice versa. Global sensitivity analysis makes use of this fact by considering main effect sensitivity indices, which can be written as

$$S_i = \frac{\text{var}(\mathbb{E}[Y|X_i])}{\text{var}(Y)}, \quad (3.17)$$

which by the definition of total variance, is related to the expected value of the variance of an output through

$$\mathbb{E}[\text{var}(Y|X_i)] = \text{var}(Y) - S_i \text{var}(Y). \quad (3.18)$$

The relation given in Equation 3.18 is the foundation of distributional sensitivity analysis.

As noted previously, the key contribution of distributional sensitivity analysis is the extension of global sensitivity analysis to situations where the amount of variance that can be reduced for a given factor is considered to be a random variable rather than assuming the variance to be completely reducible. Let  $X_i^o$  be the random variable defined by the original distribution for some factor  $i$ , and  $X_i'$  be the random variable defined by a new distribution for factor  $i$  after some further research has been done,

which have corresponding main effect sensitivity indices  $S_i^o$  and  $S_i'$  respectively. Then we can define the ratio of the variance of factor  $i$  that cannot be reduced and the total variance of the original distribution of factor  $i$  as,

$$\delta = \frac{\text{var}(X_i')}{\text{var}(X_i^o)}. \quad (3.19)$$

Assuming further research reduces the variance of factor  $i$ ,<sup>1</sup> it is clear that  $\delta \in [0, 1]$ . Since it cannot be known in advance how much variance reduction for a given factor is possible through further research, the distributional sensitivity analysis method casts  $\delta$  as a uniform random variable,  $\Delta$ , on  $[0, 1]$ , in keeping with the principle of maximum uncertainty discussed in Section 2.2.

Given that the variance of factor  $i$  that may be reduced is a random percentage,  $100(1 - \Delta)\%$ , of the total original variance of factor  $i$ , a distributional sensitivity index function can be defined as

$$\text{adjS}_i(\delta) = \frac{\text{var}(Y^o)S_i^o - \mathbb{E}[\text{var}(Y')S_i'|\Delta = \delta]}{\text{var}(Y^o)}, \quad (3.20)$$

where  $\text{adjS}_i$  is to be read as, “the adjusted main effect sensitivity index of factor  $i$ ,”  $S_i^o$  is the original main effect sensitivity index of factor  $i$ , and  $\mathbb{E}[\text{var}(Y')S_i'|\Delta = \delta]$  is the expected value of the product of the variance of the output and the main effect global sensitivity index of factor  $i$  taken over all *reasonable* distributions of factor  $i$  with  $100\delta\%$  of the variance of the original distribution for factor  $i$ . What is meant by reasonable is discussed in the following subsection.

The adjusted main effect sensitivity index function given by Equation 3.20 is interpreted as the main effect sensitivity index for factor  $i$  if it is known that only  $100(1 - \delta)\%$  of the factor’s variance can be reduced. This can be seen by noting that  $\text{var}(Y^o)S_i^o$  is the expected value of the variance of  $Y^o$  that is due to factor  $i$ , as shown in Equation 3.18, and  $\text{var}(Y')S_i'$  is the expected value of the variance of  $Y'$  that is due to factor  $i$  after  $100(1 - \delta)\%$  of factor  $i$ ’s variance has been reduced. Since there

---

<sup>1</sup>It is possible that further research could increase the variance of a factor, however, this would suggest that the original characterization of uncertainty was flawed.

are many ways to reduce the variance of factor  $i$  by  $100(1 - \delta)\%$ , the expected value of  $\text{var}(Y')S'_i$  is taken over all the reasonable distributions for which  $100(1 - \delta)\%$  has been reduced. Thus,  $\text{var}(Y^o)S_i^o - \mathbb{E}[\text{var}(Y')S'_i|\Delta = \delta]$  is amount of variance in  $Y^o$  that cannot be reduced further if factor  $i$ 's variance can only be reduced by  $100(1 - \delta)\%$ .

If it is assumed that all of the variance of a particular factor can be reduced, then  $\delta = 0$ , and for a given factor  $i$ , this means that  $\mathbb{E}[\text{var}(Y')S'_i|\Delta = 0] = 0$ , since once all of the variance of factor  $i$  has been reduced, factor  $i$  will simply become a constant, and thus,  $S'_i = 0$ . Therefore, when  $\delta = 0$ ,  $\text{adjS}_i(0) = S_i^o$ , and distributional sensitivity analysis reduces to the specific case of global sensitivity analysis. However, as noted previously, since it is not likely known what value  $\delta$  will take prior to further research on a given factor, in general distributional sensitivity analysis,  $\delta$  is considered to be a uniform random variable,  $\Delta$ , on the interval  $[0, 1]$ . The expected value of  $\text{adjS}_i(\Delta)$  can thus be taken to give an *average adjusted main effect sensitivity index* (AAS), as shown in Equation 3.21 for some factor  $i$ ,

$$\text{AAS}_i = \mathbb{E}[\text{adjS}_i(\Delta)]. \quad (3.21)$$

The average adjusted main effect sensitivity index for each factor in a model is then an index that can be used to quantitatively rank factors based on the average amount of output variance that can be reduced when further research is done on a particular factor.

### 3.3.2 Defining reasonable distributions

Several times in the discussion of the development of distributional sensitivity analysis it was mentioned that *reasonable* new factor distributions, which represent the result of further research on a factor, be used in the estimation of adjusted main effect sensitivity indices. This is because given some initial distribution for a factor and some  $\delta$ , there will generally not be a single unique new distribution with  $100\delta\%$  of the variance of the original factor distribution. For example, if a factor has an original distribution that is say uniform on the interval  $[0, 1]$ , and  $\delta = 0.5$ , there are an

infinite number of new distributions, such as  $U[0, \sqrt{2}/2]$ ,  $U[1 - \sqrt{2}/2, 1]$ ,  $U[\sqrt{2}/4, 1 - \sqrt{2}/4]$ , etc., that all have variances equal to  $(\delta)$  times the original variance. The new distributions could also be from a different family of distributions, such as triangular. Therefore, a set of reasonable distributions with  $100\delta\%$  of the variance of any given original distribution must be defined. This is done for the uniform, triangular, and normal distribution families in the following paragraphs. In each case, it is assumed that future research will only impact the given constraints regarding uncertainty information that were used to assign distributions based on the maximum uncertainty principle. Thus, the impact of future research will be studied only through changes in the parameters of a given family of distributions, not of the family of distributions itself. However, if the distribution family of a given factor was expected to change through further research (e.g. from an original uniform distribution to a distribution in the triangular family), then reasonable distributions from the new family, given that the original distribution was from another family, could be defined.

Consider an arbitrary uniform distribution,  $U[a, b]$ , as shown in Figure 3-2, where  $u(x)$  is the probability density function for some factor,  $x$ ,  $a$  and  $b$  are the endpoints of the interval where  $u(x) > 0$ , and  $h = 1/(b - a)$ . The variance of this distribution

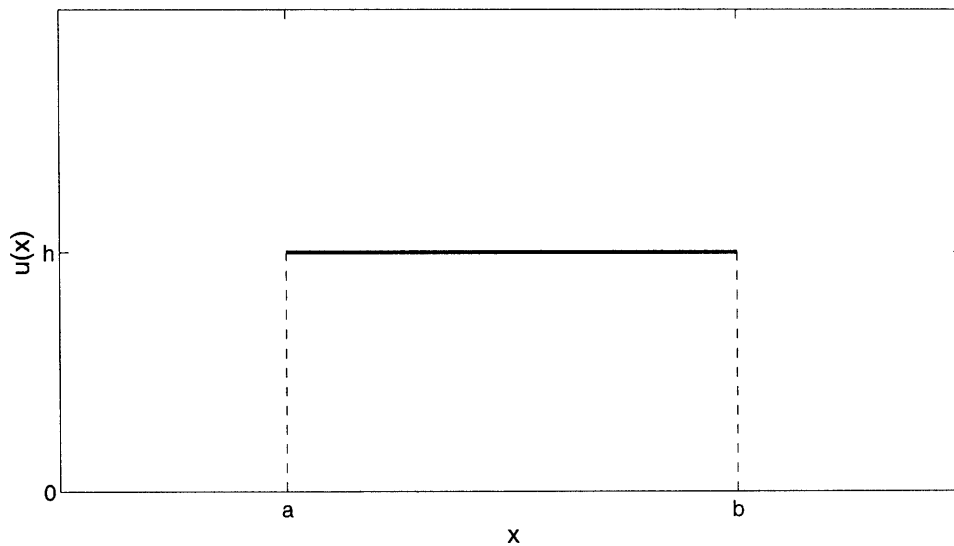


Figure 3-2: Example of a uniform probability density function on  $[a, b]$ .

is given as

$$\text{var}(X) = \frac{(b - a)^2}{12}. \quad (3.22)$$

Thus,  $\delta$  for this family of distributions can be written as

$$\delta = \left( \frac{b' - a'}{b^o - a^o} \right)^2, \quad (3.23)$$

where  $a'$  and  $b'$  are the endpoints of a new distribution and  $a^o$  and  $b^o$  are the endpoints of the original distribution. In this case, a given  $\delta$  forces all new uniform distributions that satisfy Equation 3.23 to be intervals of the same width, which is  $\delta^{1/2}(b^o - a^o)$ . A reasonable method for sampling from the set of intervals on  $[a^o, b^o]$  with width  $\delta^{1/2}(b^o - a^o)$ , is as follows:

1. Sample  $\delta$  from a uniform distribution on the interval  $[0, 1]$ .
2. Sample  $b'$  from a uniform distribution on the interval  $[a^o + \delta^{1/2}(b^o - a^o), b^o]$ .
3. Let  $a' = b' - \delta^{1/2}(b^o - a^o)$ .

This method of sampling ensures the new parameters,  $a'$  and  $b'$ , for a given  $\delta$ , will be such that  $a' \sim U[a^o, b^o - \delta^{1/2}(b^o - a^o)]$ , and  $b' \sim U[a^o + \delta^{1/2}(b^o - a^o), b^o]$ . This ensures that the set of possible uniform distributions that satisfies Equation 3.23 is sampled from uniformly. This can be seen by considering that for a given  $\delta$ , the set of possible values for the endpoint  $a'$ , is  $[a^o, b^o - \delta^{1/2}(b^o - a^o)]$  and the set of possible values for the endpoint  $b'$ , is  $[a^o + \delta^{1/2}(b^o - a^o), b^o]$ , which have both been sampled from uniformly.

Consider now an arbitrary triangular distribution,  $T(a, b, c)$ , as shown in Figure 3-3, where  $t(x)$  is the probability density function for some factor,  $x$ ,  $a$  and  $b$  are the minimum and maximum values factor  $x$  can take,  $c$  is the mode of the distribution, and  $h = 2/(b - a)$ . The variance of this distribution is given as

$$\text{var}(X) = \frac{a^2 + b^2 + c^2 - ab - ac - bc}{18}. \quad (3.24)$$

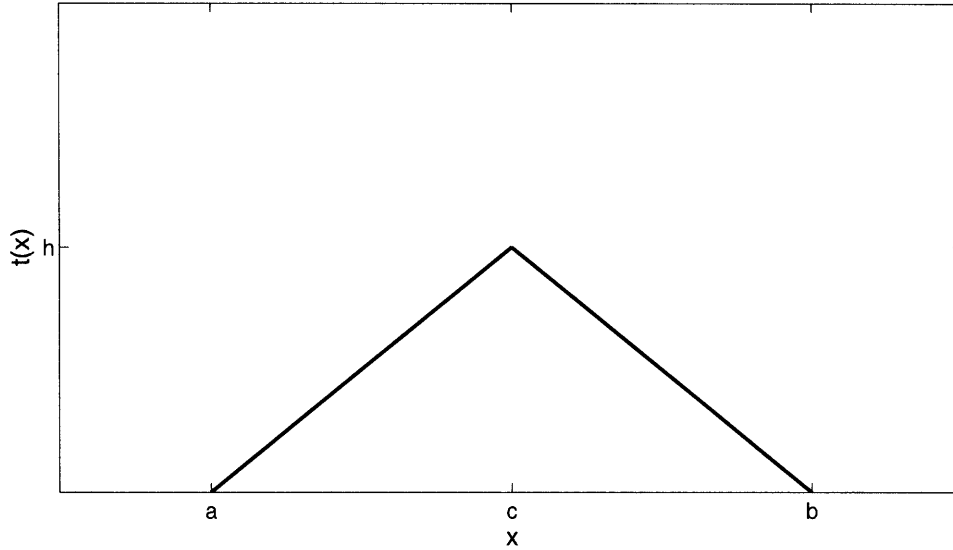


Figure 3-3: Example of a triangular probability density function with minimum value = a, maximum value = b, and mode = c.

Thus,  $\delta$  for the triangular family of distributions can be written as

$$\delta = \frac{a'^2 + b'^2 + c'^2 - a'b' - a'c' - b'c'}{a^{o2} + b^{o2} + c^{o2} - a^o b^o - a^o c^o - b^o c^o}, \quad (3.25)$$

where  $a'$ ,  $b'$ , and  $c'$  are the parameters of a new distribution, and  $a^o$ ,  $b^o$ , and  $c^o$  are the parameters of the original distribution. More is known regarding the uncertainty associated with a factor that has been assigned a triangular distribution according to the principle of maximum uncertainty than for those factors that have been assigned uniform distributions. In the case of a triangularly distributed factor, a most likely value that the factor can take exists. Here, a procedure is presented for sampling from a family of triangular distributions where the most likely value does not change, however, if it is expected that the most likely value will change, other procedures can be developed. The proposed procedure for sampling from a family of triangular distributions is as follows:

1. Set  $c' = c^o$ .
2. Sample  $\delta$  from a uniform distribution on the interval  $[0, 1]$ .
3. Sample  $u$  from a uniform distribution on the interval  $[a^o, b^o]$ .
4. If  $u \leq c'$ 
  - Set  $a' = u$ .
  - Set  $b'$  from Equation 3.25.
  - If  $b' \geq c'$  and  $b' \leq b^o$ , then accept the distribution defined as  $T(a', b', c')$ .
  - Else go back to step 3.
5. Else
  - Set  $b' = u$ .
  - Set  $a'$  from Equation 3.25.
  - If  $a' \geq a^o$  and  $a' \leq c^o$ , then accept the distribution defined as  $T(a', b', c')$ .
  - Else go back to step 3.

For the case of triangular distributions, the set of possible minimum values  $a'$  can take, and maximum values  $b'$  can take, for a given  $\delta$ , are both intervals like the case of the uniform distribution family. However, applying a procedure where the values for say  $a'$ , are sampled uniformly, will not lead to a uniform sampling of the possible  $b'$  values, as shown in Figure 3-4. Here, the original distribution is  $T(0, 2, 1)$ . The top two plots are the samples drawn uniformly from the possible values of the minimum (left) and the resulting values of the maximum (right), for  $\delta = 0.5$ . The bottom two plots are samples drawn using the proposed procedure for sampling from a family of triangular distributions. In the case of the top two plots, the minimum and maximum

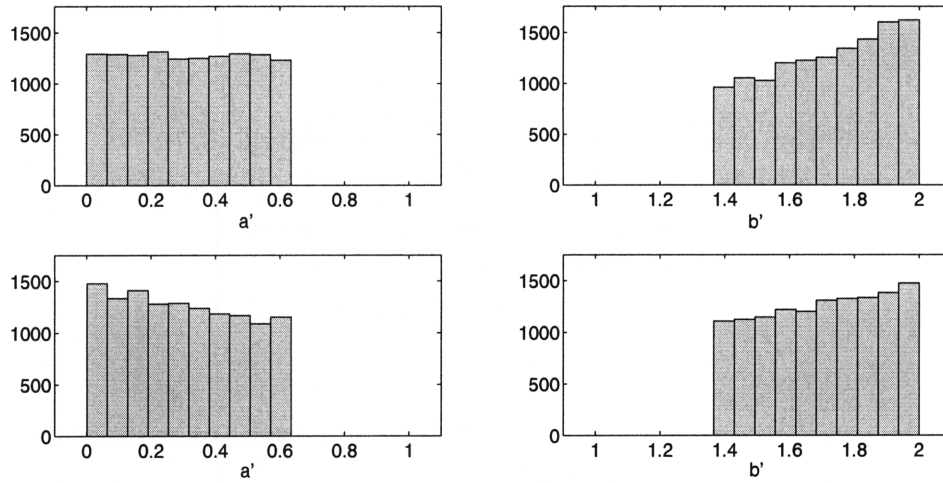


Figure 3-4: The top two plots are the samples drawn uniformly from the possible values of the minimum (left) and the resulting values of the maximum (right). The bottom two plots are the samples drawn using the proposed procedure for sampling from a family of triangular distributions.

values of the new distributions are not sampled from in the same fashion, whereas in the case of the proposed procedure, the empirical evidence suggests that the endpoints of the new distributions are sampled from in the same manner.

Finally, consider a normal distribution,  $\mathcal{N}(\mu, \sigma^2)$ , shown in Figure 3-5, where  $\mu$  is the mean and  $\sigma^2$  is the variance of the distribution. For the normal family of distributions,  $\delta$  is given as

$$\delta = \frac{\sigma'^2}{\sigma^2}, \quad (3.26)$$

where  $\sigma'^2$  is the variance of a new distribution and  $\sigma^2$  is the original variance. Here a procedure is presented where the mean value of the original distribution is also the mean value of any new distributions after further research has been undertaken. However, if the mean value is expected to change, other procedures can be developed to take that into account. Given that here the mean does not change,  $\delta$  uniquely defines new distributions with the property given by Equation 3.26, where given some  $\delta$ , there is only one possible new distribution, which is  $\mathcal{N}(\mu, \delta\sigma^2)$ . The proposed procedure for sampling from a family of normal distributions is simply:



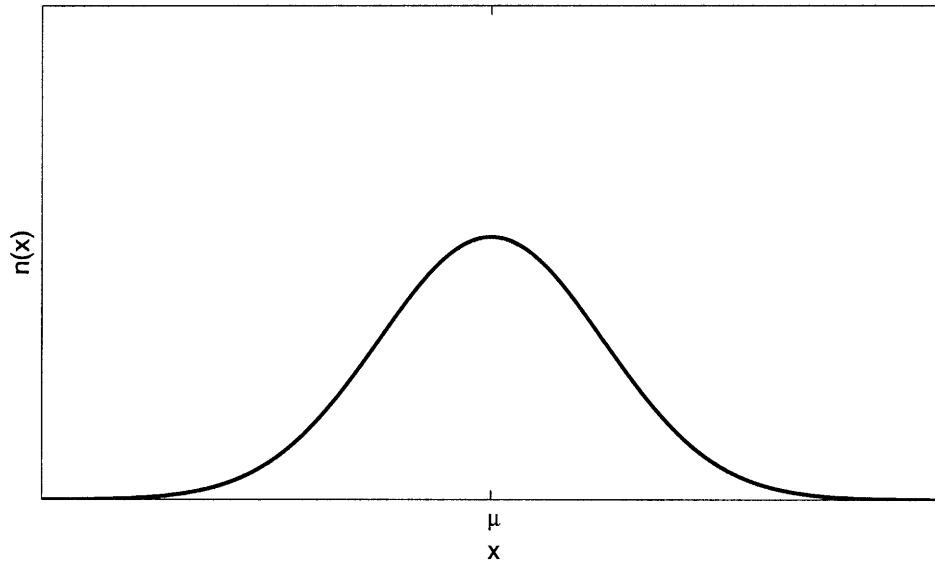


Figure 3-5: Example of a normal probability density function with mean =  $\mu$  and variance =  $\sigma^2$ .

1. Set  $\mu' = \mu^o$ .
2. Sample  $\delta$  from a uniform distribution on the interval  $[0, 1]$ .
3. Set  $\sigma'^2 = \delta\sigma^2$ .

### 3.3.3 Acceptance/rejection sampling

Clearly the evaluation of Equation 3.20 and subsequently of Equation 3.21 requires consideration of a large number of different distributions for each factor. If a global sensitivity analysis is carried out for each new distribution for each factor, the computational expense would be massive and distributional sensitivity analysis would likely be too costly to ever carry out. However, if a global sensitivity analysis with the original distributions for each factor is completed, a method known as acceptance/rejection sampling can be used to complete a distributional sensitivity analysis without any further model evaluations.

Acceptance/rejection sampling is a method for generating samples from a desired distribution by sampling from a different distribution. The method is based on random sampling of the original distribution and is thus well-established only for brute force pseudorandom sampling, though work is being done to extend the method to other sampling strategies [54]. Following Degroot and Schervish [55], let  $g(z)$  be a probability density function of a desired distribution for some random variable,  $Z$ . Let  $p(x)$  be some other probability density function for a random variable,  $X$ , with the property that there exists a constant,  $k$ , such that  $kp(x) \geq g(x)$  for all  $x$ . The acceptance/rejection method can then be used to generate samples from  $g(z)$  as follows:

1. Draw a sample,  $x$ , from  $p$ .
2. Draw a sample,  $u$ , from a uniform random variable on  $[0, 1]$ .
3. If  $\frac{g(x)}{p(x)} \geq ku$ , let  $z = x$ .
4. Else, discard  $x$  and  $u$  and return to step 1.

The process may then be repeated until the desired distribution has been sampled from sufficiently.

Figures 3-6 through 3-8 show examples of uniform, triangular, and normal distributions drawn using acceptance/rejection sampling and using the methods discussed in Section 3.3.2 to establish the parameters of the distributions. The original distributions were  $U[0, 1]$ ,  $T(0, 2, 1)$ , and  $\mathcal{N}(0, 1)$  for the uniform, triangular, and normal distributions respectively. For each new distribution,  $\delta$  was set at 0.5. As can be seen for the case of the uniform distribution in Figure 3-6, there is more than one distribution that can be drawn with half the variance of the original distribution. This is also the case for the triangular distribution as shown in Figure 3-7. This implies that the adjusted main effect sensitivity index,  $\text{adjS}_i(\delta)$ , for a given  $\delta$ , will be a random variable. Thus, in the case of the uniform and triangular distribution, for each  $\delta$ , the

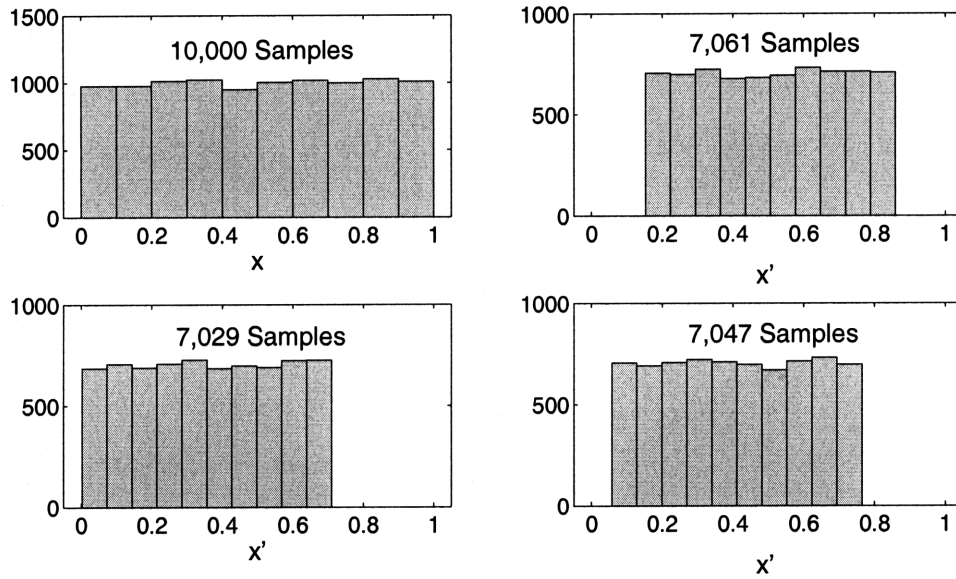


Figure 3-6: Example histograms that result from randomly sampling from an original uniform random variable on  $[0,1]$  (top left), and the use of acceptance/rejection sampling on the original samples with  $\delta = 0.5$  (top right, lower left, lower right).

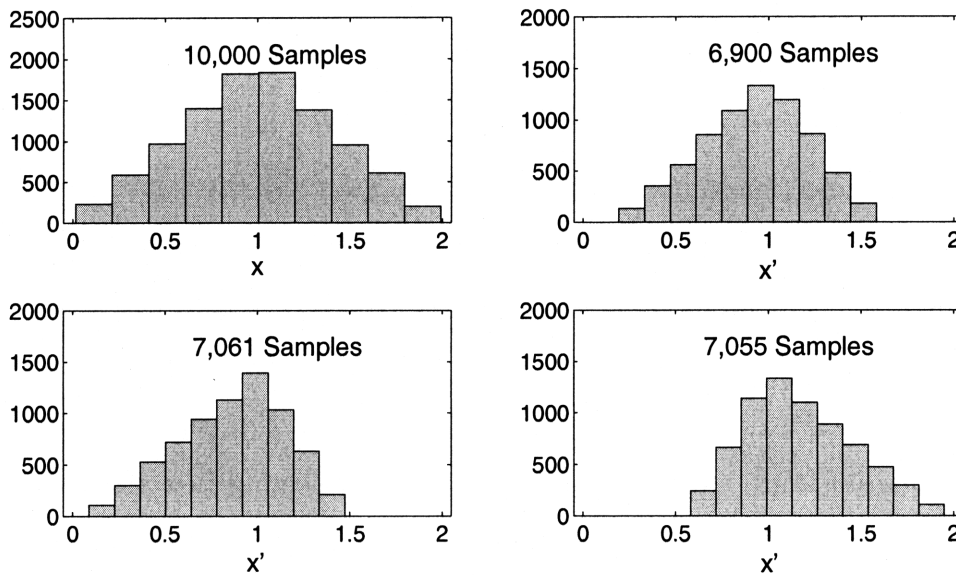


Figure 3-7: Example histograms that result from randomly sampling from an original triangular random variable with parameters  $(0,2,1)$  (top left), and the use of acceptance/rejection sampling on the original samples with  $\delta = 0.5$  (top right, lower left, lower right).

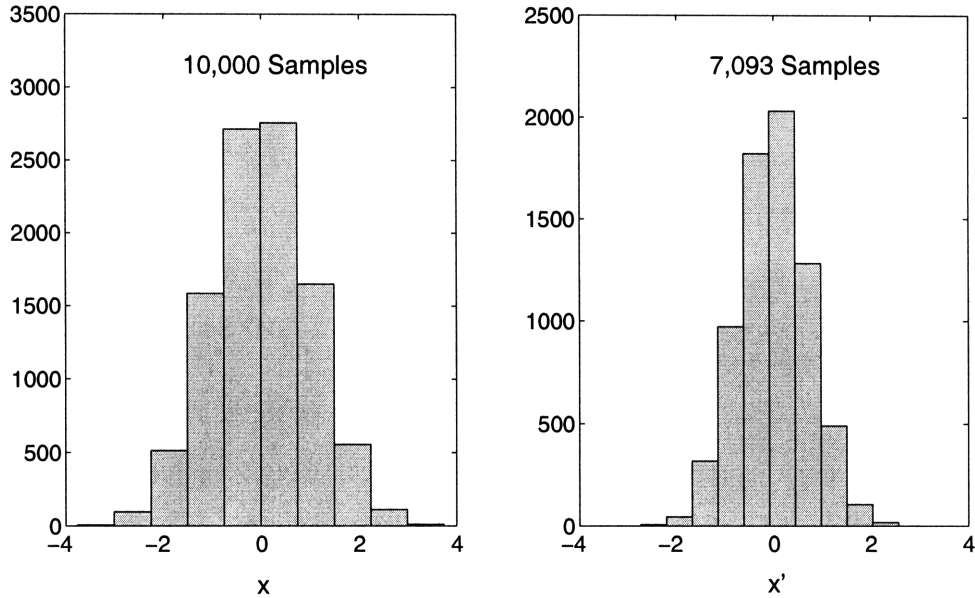


Figure 3-8: Example histograms that result from randomly sampling from an original normal random variable with mean = 0 and variance = 1 (left), and the use of acceptance/rejection sampling on the original samples with  $\delta = 0.5$  (right).

adjusted main effect sensitivity index must be sampled from several times to arrive at a suitable estimate of the mean of the index for each  $\delta$ .

Acceptance/rejection sampling can be employed to reuse the results from a global sensitivity analysis in a distributional sensitivity analysis as follows. Consider some factor,  $i$ , with an original distribution defined as  $T(a^o, b^o, c^o)$  for which a global sensitivity analysis has been conducted. Thus, we have the model evaluations corresponding to  $f([x_1^m, \dots, x_i^m, \dots, x_n^m]^T)$  and  $f([\tilde{x}_1^m, \dots, x_i^m, \dots, \tilde{x}_n^m]^T)$ , where  $m = 1, \dots, N$ , as given in Equation 3.13. Assume the procedure for selecting new triangular distributions given in Section 3.3.2 selects new parameters  $a'$ ,  $b'$ , and  $c'$ , so that we have a new distribution defined as  $T(a', b', c')$ , for which we would like to estimate the main effect sensitivity index,  $S'_i$ , and the new variance of the output,  $\text{var}(Y')$ , for use in the estimation of an adjusted main effect sensitivity index for factor  $i$  given by Equation 3.20. The estimation of these quantities using the function evaluations from global sensitivity analysis can be achieved with the following algorithm.

1. Set  $m = 1, j = 1$
2. Set  $k = \frac{b^o - a^o}{b' - a'}$
3. Draw a sample,  $u$ , from a uniform random variable on  $[0,1]$
4. If  $\frac{g(x^m)}{p(x^m)} \geq ku$ 
  - Let  $h([x_1^j, \dots, x_i^j, \dots, x_n^j]^T) = f([x_1^m, \dots, x_i^m, \dots, x_n^m]^T)$
  - Let  $h([\tilde{x}_1^j, \dots, x_i^j, \dots, \tilde{x}_n^j]^T) = f([\tilde{x}_1^m, \dots, x_i^m, \dots, \tilde{x}_n^m]^T)$
  - $j = j + 1, m = m + 1$
5. Else  $m = m + 1$
6. If  $m < N$  go to 3
7. Else
  - $M = j$
  - $\hat{h}_o = \frac{1}{M} \sum_{j=1}^M h([x_1^j, \dots, x_i^j, \dots, x_n^j]^T)$
  - $\hat{D}' = \frac{1}{M} \sum_{j=1}^M h([x_1^j, \dots, x_i^j, \dots, x_n^j]^T)^2 - \hat{h}_o^2$
  - $\hat{D}'_i = \frac{1}{M} \sum_{j=1}^M h([x_1^j, \dots, x_i^j, \dots, x_n^j]^T) \cdot h([\tilde{x}_1^j, \dots, x_i^j, \dots, \tilde{x}_n^j]^T) - \hat{h}_o^2$
  - $\hat{S}'_i = \frac{\hat{D}'_i}{\hat{D}'}$
  - $\text{var}(Y') = \hat{D}'$

### 3.3.4 Example of global and distributional sensitivity analysis

The following example reveals the benefit of using distributional sensitivity analysis in favor of global sensitivity analysis results for factor prioritization. Consider a model

given by

$$f(X_1, X_2, X_3) = \frac{1}{10}\exp(X_1) + 20\exp(-X_2) + 11X_3, \quad (3.27)$$

where  $X_1 \sim T(0, 6, 1/2)$ ,  $X_2 \sim T(0, 6, 1/2)$ , and  $X_3 \sim T(0, 2, 1)$ . Distributional sensitivity analysis was conducted on each factor of the model as described in the preceding sections. Figure 3-9 presents the adjusted main effect sensitivity indices of each factor for values of  $\delta = 0.0, 0.1, 0.2, \dots, 1.0$ . The figure shows clearly that the

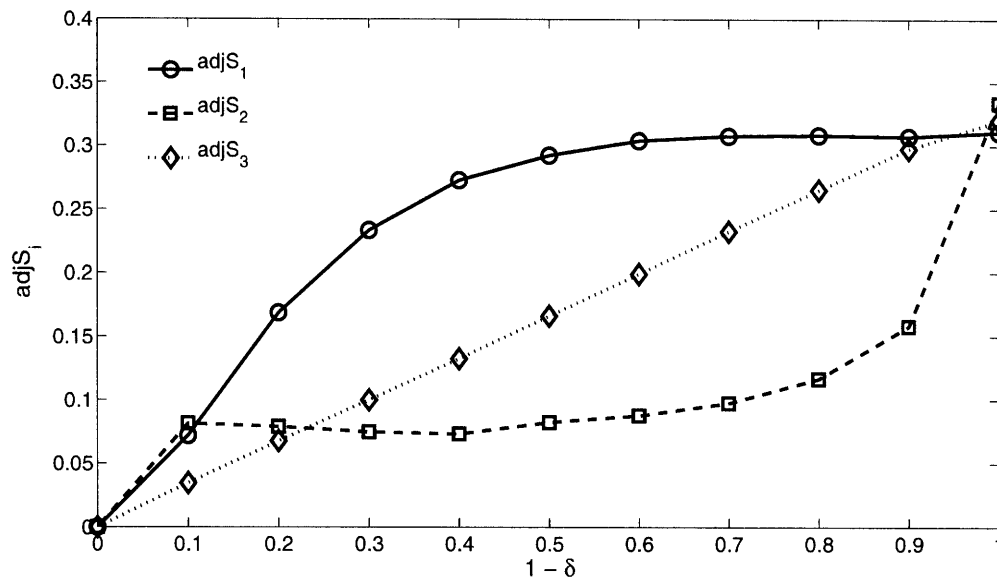


Figure 3-9: Adjusted main effect sensitivity indices of each factor for values of  $\delta = 0.0, 0.1, 0.2, \dots, 1.0$ . The adjusted main effect sensitivity indices of each factor for  $\delta = 0$  (rightmost points), are the main effect sensitivity indices of each factor as computed by global sensitivity analysis.

factors that should be considered for further research to reduce output variance depend on the amount of variance that is assumed reducible for each factor. Figure 3-10 compares the main effect sensitivity indices estimated from global sensitivity analysis for each factor with the average adjusted main effect sensitivity indices estimated via distributional sensitivity analysis. The global sensitivity results suggest that the ranking for factor prioritization be factor 2, followed by factor 3, and then factor 1, though all of the indices are close to one another. The distributional sensitivity results however, suggest that the ranking for factor prioritization be factor 1, followed

by factor 3, and then factor 2, with clear differences in the magnitudes of the three indices. Thus, in the case of this example model, assuming the variance of a given factor can be reduced to zero through further research, as is done in the global sensitivity analysis factor prioritization setting, leads to a completely different conclusion regarding which factors should be researched than distributional sensitivity analysis, which assumes the amount of variance that can be reduced for a given factor is a random variable that is uniformly distributed over the range of reducing none of the variance to all of the variance of the particular factor. Since, as previously stated, the notion that all of the variance of a given factor can be reduced through further research is optimistic, the distributional sensitivity analysis results are considered more reliable and are recommended for use in favor of global sensitivity analysis for factor prioritization. A comparison of both methods for use in the factor prioritization for a real world application is presented in Chapter 5.

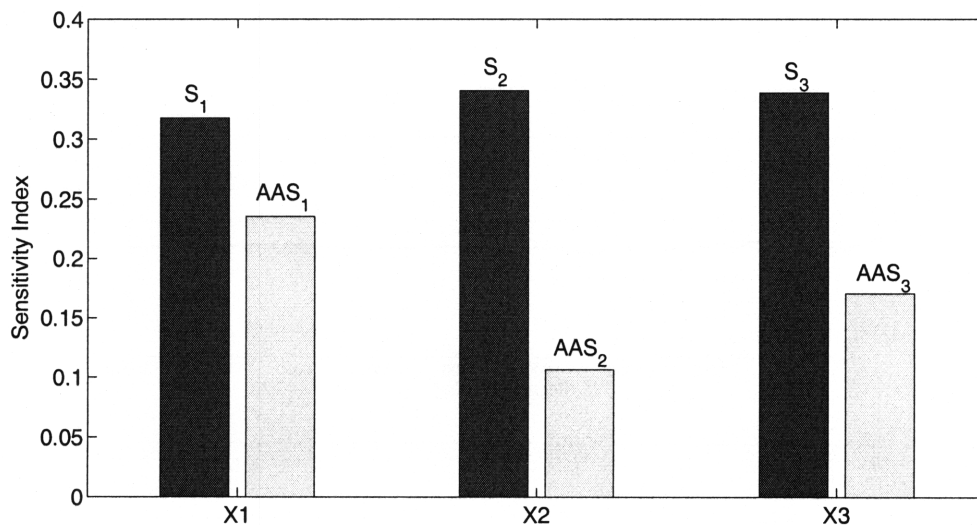


Figure 3-10: A comparison of the main effect sensitivity indices estimated from global sensitivity analysis to the average adjusted main effect sensitivity indices estimated from distributional sensitivity analysis.

### 3.4 Convergence of Uncertainty and Sensitivity Analyses

It is important to note that estimates of statistical quantities described in this Chapter, such as the total and main effect sensitivity indices, the adjusted main effect sensitivity indices, and the average adjusted sensitivity indices, as well as estimates of output means and variances discussed in Chapter 2, via Monte Carlo simulation, can take many thousands of model evaluations. As more and more model evaluations are conducted, the statistical quantities converge to their true values as guaranteed by the law of large numbers. This convergence can be studied graphically to determine whether or not more function evaluations are required for each statistical quantity being estimated. For the model studied in the previous subsection, Figure 3-11 presents a graphical look at the estimates of the mean and variance as the number of iterations increases. As can be seen from the figure, the variance estimation shows more variability than the mean estimation. Given that both global and

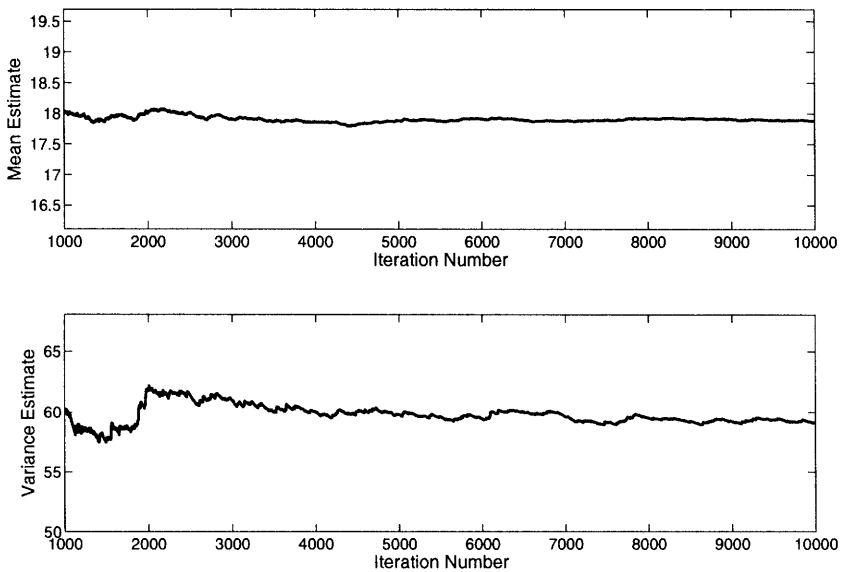


Figure 3-11: Estimates of the mean and variance of the test model output as the number of iterations increases.



distributional sensitivity analyses focus on estimates of variance, these methods tend to be more computationally expensive than methods such as the mean-value method that focus on mean estimates. Figure 3-12 gives the estimates of the total, main, and average adjusted main effect sensitivity indices as the number of iterations used in the global sensitivity analysis increases. The convergence of individual adjusted

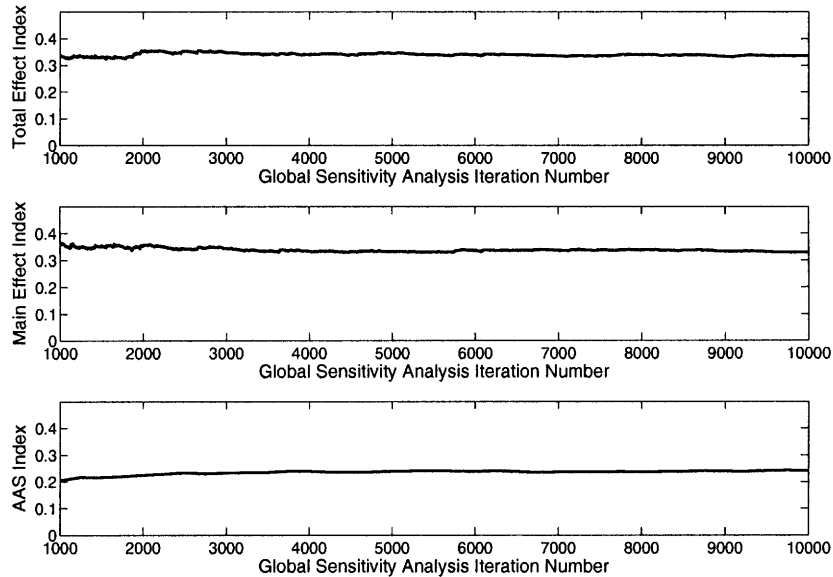


Figure 3-12: Estimates of the total, main, and average adjusted main effect sensitivity indices of factor  $X_1$  as the number of iterations used for global sensitivity analysis increases.

main effect sensitivity indices is shown in Figure 3-13 for values of  $\delta$  equal to 0.9 (top), 0.6 (middle), and 0.3 (bottom). To determine how many distributions should be considered for each  $\delta$  value in the calculation of the adjusted main effect sensitivity indices, bootstrap confidence intervals can be constructed for each  $\text{adjS}_i(\delta)$  as shown in Ref. [56]. The results for factor  $X_1$  of the test model using 50 reasonable distributions for each  $\delta$  are shown in Figure 3-14.

Figures such as those shown in this subsection, can be used adaptively, if necessary, to decide when to stop sampling. For a situation where the results of a global sensitivity analysis have converged sufficiently but the results of the distributional sensitivity analysis have not, it is recommended that more samples be added to the

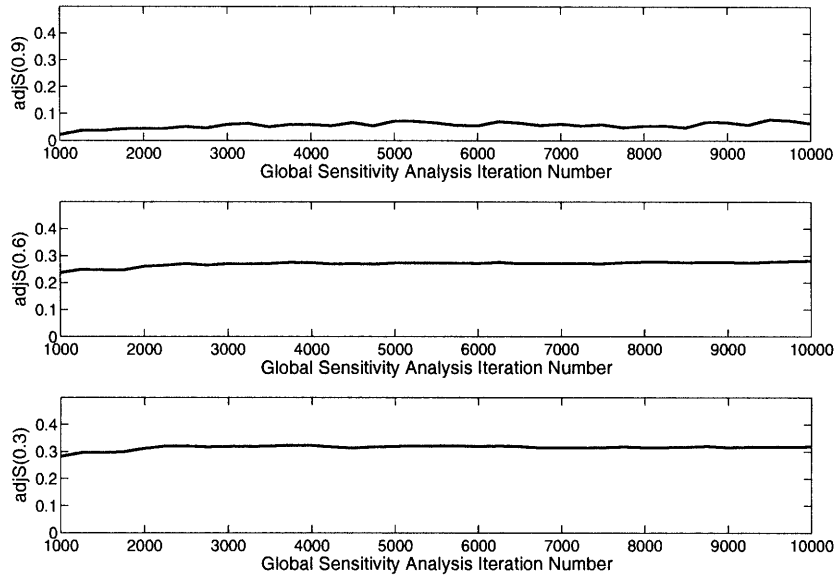


Figure 3-13: Estimates of the adjusted main effect sensitivity indices for  $\delta = 0.9$  (top),  $\delta = 0.6$  (middle), and  $\delta = 0.3$  (bottom), for factor  $X_1$  as the number of iterations used for global sensitivity analysis increases.

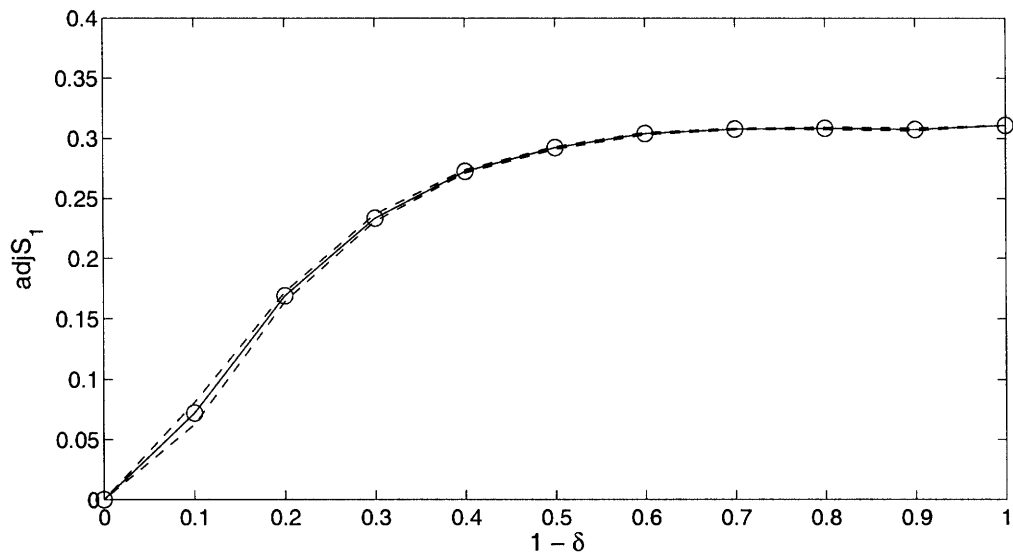


Figure 3-14: Bootstrap confidence intervals for the adjusted main effect sensitivity indices of factor  $X_1$  using 50 reasonable distributions for each value of  $\delta$ .

global sensitivity analysis, which will lead to an increase in the number of samples used in the distributional sensitivity analysis. For situations where global or distribu-

tional sensitivity analyses require a large number of samples to converge sufficiently, surrogate models can be developed and implemented, as presented in the following chapter, to address concerns regarding computational expense.



# Chapter 4

## Surrogate Modeling for Uncertainty Assessment

As discussed in the previous chapter, estimates of statistical quantities of interest in uncertainty and sensitivity analyses can require many thousands of model evaluations. Often, complex models have long computation times, which implies that a large number of model evaluations could present an intractable computational burden. In such situations, surrogate models that provide substantial computational speedups are crucial to the process of uncertainty assessment. Further, knowledge of the quantitative impacts on the analyses associated with exercising a surrogate in place of a full model is essential to producing defensible claims in the context of decision-making.

This chapter presents a systematic method to reduce the complexity and computational cost of a complex model designed to estimate global emissions from aviation, in such a way that factor uncertainty may still be quantified and analyzed. Section 4.1 presents background material on surrogate modeling methodologies. Section 4.2 describes the structure of the specific case considered here—the Aircraft Emissions Model (AEM) within AEDT. The specific methodology developed here is described in Section 4.3, and focuses on the creation of a hierarchical surrogate model to represent the complex system. Section 4.4 presents the results of applying the methodology to the AEM. The work presented in this chapter focuses on the

systematic development of a surrogate model for use in Steps 5 and 6 (conducting uncertainty and sensitivity analyses) of the general approach discussed in Chapter 2 for a specific real-world model. An example application of the entire general approach to a real-world model will be presented in Chapter 5.

## 4.1 Surrogate Modeling Background

Surrogate models can be categorized into three different classes: data-fit models, reduced-order models, and hierarchical models [57]. Data-fit models are generated using interpolation or regression of simulation data from the input/output relationships in the high-fidelity model [58]. The primary challenge in adopting this surrogate modeling strategy for large-scale complex system models is the “curse of dimensionality” when the number of inputs to a model is large and design of experiment techniques must be applied with care in order to balance the computational cost of the required simulations with coverage of the input space. Reduced-order models are typically constructed for systems described by partial differential equations or large sets of ordinary differential equations [59]. Derivation of reduced-order models relies on the knowledge of the governing equations and are thus in general not suitable for systems for which the governing equations are unknown or empirically based. Hierarchical surrogate models, also known as variable fidelity models, employ simplified mathematical models such as coarser grids in finite element models [60] and models with simplified physics [44, 61].

The application of a particular surrogate modeling strategy depends both on what computational tasks are to be performed, and on the underlying structure of the model. In some cases, nothing will be known about a given model, and strategies that perform better in black-box situations, such as data-fit methods, should be used. In other cases, everything will be known about the governing equations of a given model, and reduced-order models can be derived using projection-based approaches. For the AEM, described in Section 4.2.1, the structure of the governing equations of the model is known, but the input space is too large to use a projection-based model

reduction approach or a data-fit method. However, the model structure is such that a hierarchical surrogate modeling strategy can be employed and, as described in Section 4.3, further exploited to provide quantified confidence intervals on surrogate predictions.

## 4.2 Application

As noted in Chapter 1, the real-world application that has motivated this work is the FAA tools-suite, which is shown in Figure 1-3. The scale and complexity of analyses run with this set of tools is immense; for example, a single simulation of a one-year analysis involves over thirty million flight operations with 350 aircraft types and thousands of factors. Thus, uncertainty and sensitivity analyses for some of the models within the tools-suite are computationally infeasible. Here, a systematic method to reduce the complexity and computational cost of the AEM is developed in such a way that factor uncertainty may still be assessed.

### 4.2.1 Aircraft emissions model

The AEM is used to calculate emissions inventories of such pollutants as CO<sub>2</sub>, CO, NO<sub>x</sub>, SO<sub>x</sub>, and many others. The calculation is done on an operation-by-operation basis, and the emissions computed for each operation in a given scenario are then aggregated to produce an emissions inventory. An operation is in turn simulated on a flight segment-by-segment basis as shown in Figure 4-1, where emissions are calculated for each segment of the operation and then aggregated to produce the total emissions of the operation.

The six AEM factors considered are shown in Table 4.1, where each factor is defined for each segment of each operation. We consider the emissions resulting from a total of  $N_o$  operations, each consisting of  $N_s$  flight segments. Thus, the total number of factors is given by  $n = 6N_oN_s$ . Table 4.1 also shows the probability density functions that are defined for each factor on a segment-by-segment basis. These density functions were arrived at through previous studies and expert opinions.

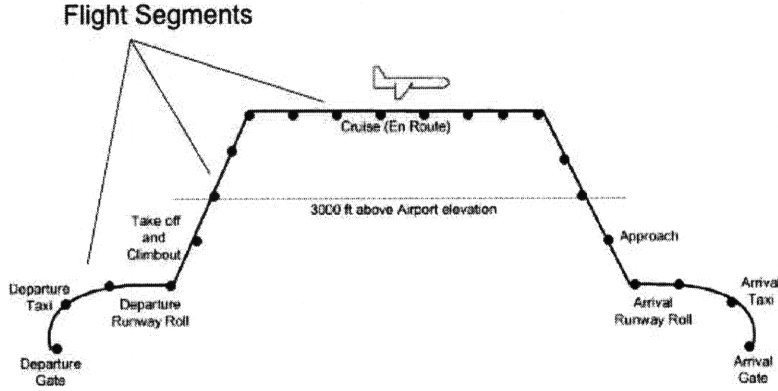


Figure 4-1: AEM segmentation of an aircraft operation. Adapted from Ref. [2].

Table 4.1: AEM factors and their probability density functions. All factors are applied as multipliers to nominal factor values.

Input Variable	Input Quantity	Distribution Type	Defining Values
$x_1 := q$	Fuel Burn	Uniform	[0.95, 1.05]
$x_2 := r$	Temperature	Triangular	[0.89, 1.00, 1.11]
$x_3 := s$	Pressure	Triangular	[0.97, 1.00, 1.03]
$x_4 := t$	Relative Humidity	Triangular	[0.82, 1.00, 1.17]
$x_5 := u$	Fuel Flow	Uniform	[0.95, 1.05]
$x_6 := v$	REINO <sub>x</sub>	Triangular	[0.76, 1.00, 1.24]

The samples from the density functions are applied as multipliers to default values of the various factors that are specific to aircraft type, engine type, and geographic location. For triangular distributions, the defining values are the minimum, mode, and maximum values. For uniform distributions, the defining values are the minimum and the maximum.

The outputs of the AEM (global emissions of NO<sub>x</sub>, CO, CO<sub>2</sub>, etc.) are all computed in a similar manner, thus the modeling methodology is developed here only for the NO<sub>x</sub> output. Other outputs are treated in an analogous way. The NO<sub>x</sub> produced for operation  $l$ ,  $y_l$ , is calculated as

$$y_l = \sum_{k=1}^{N_s} q_{kl} g_{kl}(r_{kl}, s_{kl}, t_{kl}, u_{kl}, v_{kl}), \quad (4.1)$$

where  $q_{kl}$  is the fuelburn on segment  $k$  of operation  $l$ , and  $g_{kl}(r_{kl}, s_{kl}, t_{kl}, u_{kl}, v_{kl})$  is the



emissions index of  $\text{NO}_x$  ( $\text{EINO}_x$ ) on segment  $k$  of operation  $l$ , which is calculated using Boeing Method 2 with the factors defined in Table 4.1, specifically the temperature, pressure, relative humidity, fuel flow and reference emissions index of  $\text{NO}_x$  ( $\text{REINO}_x$ ) for the given segment [62]. The total  $\text{NO}_x$  output for a set of  $N_o$  operations is then calculated as

$$y_{tot} = \sum_{l=1}^{N_o} y_l = \sum_{l=1}^{N_o} \sum_{k=1}^{N_s} q_{kl} g_{kl}(r_{kl}, s_{kl}, t_{kl}, u_{kl}, v_{kl}). \quad (4.2)$$

Since the functions within the AEM are all continuous, and each input of the AEM is a random variable, each output of the AEM is also a random variable. Thus, the total  $\text{NO}_x$  output,  $y_{tot}$ , can be thought of as a random sample from the random variable  $Y_{tot}$ , for which confidence intervals and sensitivity indices are desired.

Equation 4.2 reveals the structure of the AEM once it has been decomposed by operations. Given that computations are performed separately on each operation within the AEM, a natural representative for building a hierarchical surrogate is a single operation. Our surrogate modeling approach is thus to approximate the output of interest, total  $\text{NO}_x$  emissions, using a subset of flight operations. This is illustrated in Figure 4-2, where it can be seen that the surrogate modeling approach is based on reducing the dimension of the input space. This approach is generally applicable in many settings: for some cases (such as the AEM), reduction of the input space yields directly a hierarchical surrogate model of lower computational complexity; for other cases, the approach leads to an intermediate system of lower input dimension to which a data-fit surrogate modeling method could subsequently be applied.

The novel contribution of the methodology is to show that the structure of the AEM presented in Equation 4.2 can be exploited to use the Central Limit Theorem for calculating confidence intervals around AEM outputs computed with a hierarchical surrogate model. A similar structure exists for the ANOVA-HDMR presented in Section 3.2, which, when applied to the AEM, will also permit the use of the Central Limit Theorem for producing confidence intervals around sensitivity indices computed using the surrogate model. These methods are described in the following section.

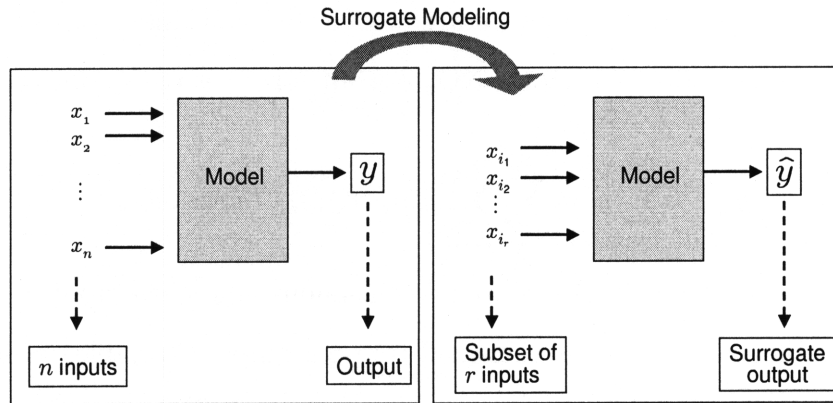


Figure 4-2: The hierarchical surrogate modeling approach achieves a reduction in computational complexity through a reduction of the input space. For the AEM, this amounts to selecting a subset of  $r$  operations, denoted by the subscripts,  $i_1, \dots, i_r$ , over which to estimate the total emissions.

### 4.3 Surrogate Modeling Methodology with Quantified Confidence Intervals

As noted previously, it is important that the effects of using a surrogate in place of a full model in uncertainty assessment be quantified. The surrogate modeling methodology developed here to perform conduct uncertainty and sensitivity analyses on the AEM is discussed in the following subsections.

#### 4.3.1 AEM surrogate for decision-making uncertainty analysis

Since the functions within the AEM are all continuous, and each factor of the AEM is a random variable, each single operation output of the AEM is also a random variable. These random variables are independent and satisfy the Lyapunov condition [63]; therefore, the Central Limit theorem may be used to completely characterize the

distribution of total  $\text{NO}_x$  emissions calculated from Equation 4.2. Here, it is noted that, owing to such circumstances as aircraft operating on similar routes, certain operations in the real-world would encounter similar environmental factors, such as temperature, pressure, and humidity, thus causing some dependence in how those factors enter into emissions estimates, which would lead to some dependence in the outputs of these operations. However, the AEM does not currently include these factor dependencies, which implies that the operation-level outputs of the AEM are completely independent of each other. Thus, the Central Limit theorem may be invoked.

According to the Central Limit theorem, the output distribution of total  $\text{NO}_x$  emissions,  $y_{tot}$ , is normally distributed since

$$y_{tot} = \sum_{l=1}^{N_o} y_l \rightarrow^d \mathcal{N}\left(\sum_{l=1}^{N_o} \mathbb{E}[y_l], \sum_{l=1}^{N_o} \sigma_{y_l}^2\right) \text{ as } N_o \rightarrow \infty, \quad (4.3)$$

where the convergence is in distribution, and  $\mathcal{N}(\alpha, \beta)$  is a normal distribution with mean  $\alpha$  and variance  $\beta$ . In Equation 4.3,  $N_o$  will not tend to infinity, however, typical analyses involving the AEM calculate emissions inventories for representative days of operations, for which  $N_o \approx 70,000$ , and one year of operations, for which  $N_o \approx 30,000,000$ . According to Ref. [64], the sample size at which  $y_{tot}$  becomes approximately normal is  $N_o \geq 30$ , thus the number of samples is much greater than required for the analyses considered here.

To estimate the distribution of  $y_{tot}$  with a surrogate model, only estimates of  $\sum_{l=1}^{N_o} \mathbb{E}[y_l]$  and  $\sum_{l=1}^{N_o} \sigma_{y_l}^2$  are required. We may estimate these quantities by noting that if we were to compute  $\mathbb{E}[y_l]$  for every operation, we could view the resulting set of expected values as representing a set of  $N_o$  samples drawn from some distribution. Thus, the expected value of  $\text{NO}_x$  emissions for some operation  $l$  can be considered as a sample from a random variable, and can be estimated using a subset of operations chosen from the full set. We denote by  $\mathcal{O}$  the subset of  $n_o$  operations chosen randomly from the full set of  $N_o$  operations. Then using the law of large numbers,  $\sum_{l=1}^{N_o} \mathbb{E}[y_l]$

is estimated as  $N_o \frac{1}{n_o} \sum_{l \in \mathcal{O}} \mathbb{E}[y_l]$ , since

$$N_o \frac{1}{n_o} \sum_{l \in \mathcal{O}} \mathbb{E}[y_l] \rightarrow N_o \mathbb{E}[\mathbb{E}[y_l]] = N_o \frac{1}{N_o} \sum_{l=1}^{N_o} \mathbb{E}[y_l] = \sum_{l=1}^{N_o} \mathbb{E}[y_l] \text{ as } n_o \rightarrow N_o. \quad (4.4)$$

For the sum of the variances of the operational level  $\text{NO}_x$  emissions in Equation 4.3, a similar method is followed to derive an analogous expression for the variance estimate of  $y_{tot}$ . Thus, the surrogate model estimate of the total  $\text{NO}_x$  output distribution using the subset  $\mathcal{O}$  of  $n_o$  operations to represent the full  $N_o$  operations is given by

$$\hat{y}_{tot} \sim \mathcal{N} \left( \frac{N_o}{n_o} \sum_{l \in \mathcal{O}} \mathbb{E}[y_l], \frac{N_o}{n_o} \sum_{l \in \mathcal{O}} \sigma_{y_l}^2 \right), \quad (4.5)$$

where  $\hat{y}_{tot}$  is a random variable that is an estimate of the random variable  $y_{tot}$ .

In Equation 4.5, the terms  $\frac{1}{n_o} \sum_{l \in \mathcal{O}} \mathbb{E}[y_l]$  and  $\frac{1}{n_o} \sum_{l \in \mathcal{O}} \sigma_{y_l}^2$ , are sample means of the distributions of expected values of  $y_l$  and of the variances of the  $y_l$ , respectively. According to the Central Limit theorem, these sample means have the following normal distributions:

$$\begin{aligned} \frac{1}{n_o} \sum_{l \in \mathcal{O}} \mathbb{E}[y_l] &\sim \mathcal{N}(\mu_{\mathbb{E}}, \frac{N_o - n_o}{N_o - 1} \sigma_{\mathbb{E}}^2 / n_o), \\ \frac{1}{n_o} \sum_{l \in \mathcal{O}} \sigma_{y_l}^2 &\sim \mathcal{N}(\mu_{\sigma^2}, \frac{N_o - n_o}{N_o - 1} \sigma_{\sigma^2}^2 / n_o), \end{aligned} \quad (4.6)$$

where  $\mu_{\mathbb{E}}$  is the expected value of the distribution of expected values of the  $y_l$ ,  $\sigma_{\mathbb{E}}^2$  is the variance of the distribution of expected values of  $y_l$ ,  $\mu_{\sigma^2}$  is the expected value of the distribution of variances of the  $y_l$ , and  $\sigma_{\sigma^2}^2$  is the variance of the distribution of variances of the  $y_l$ . The  $\frac{N_o - n_o}{N_o - 1}$  terms are finite population correction factors that must be applied since  $N_o$  is finite and the sampling of  $n_o$  operations from  $N_o$  total operations is done without replacement [65].

As noted in Section 2.3, a key outcome of an uncertainty analysis intended to support decision-making is the ability to compare such quantities as output means and variances. These quantities cannot be computed exactly using a surrogate model;

however, confidence intervals for these quantities can be rigorously computed since, as shown in the analysis above, the parameters are normally distributed. The confidence intervals for the mean and variance of total  $\text{NO}_x$  emissions can be constructed from

$$\frac{N_o}{n_o} \sum_{l \in \mathcal{O}} \mathbb{E}[y_l] - z_{\alpha/2} \sqrt{\frac{N_o - n_o}{N_o - 1} \frac{N_o^2}{n_o} \sigma_{\mathbb{E}}^2} < \mathbb{E}[y_{tot}] < \frac{N_o}{n_o} \sum_{l \in \mathcal{O}} \mathbb{E}[y_l] + z_{\alpha/2} \sqrt{\frac{N_o - n_o}{N_o - 1} \frac{N_o^2}{n_o} \sigma_{\mathbb{E}}^2} \quad (4.7)$$

and

$$\frac{N_o}{n_o} \sum_{l \in \mathcal{O}} \sigma_{y_l}^2 - z_{\alpha/2} \sqrt{\frac{N_o - n_o}{N_o - 1} \frac{N_o^2}{n_o} \sigma_{\sigma^2}^2} < \text{var}(y_{tot}) < \frac{N_o}{n_o} \sum_{l \in \mathcal{O}} \sigma_{y_l}^2 + z_{\alpha/2} \sqrt{\frac{N_o - n_o}{N_o - 1} \frac{N_o^2}{n_o} \sigma_{\sigma^2}^2}, \quad (4.8)$$

where  $z_{\alpha/2}$  is the value of the inverse cumulative distribution function of a standard normal random variable evaluated at  $(1 - \alpha/2)$ , where  $\alpha$  sets the level of confidence [52]. A typical value of  $z_{\alpha/2}$  is 1.96, which corresponds to a 95% confidence interval. In practice, constructing these confidence intervals requires estimating the variance of the distribution of the expected values of the  $y_l$ ,  $\sigma_{\mathbb{E}}^2$ , and the variance of the distribution of the variances of the  $y_l$ ,  $\sigma_{\sigma^2}^2$ . We estimate these parameters using the sample variance for each, which are calculated from

$$\hat{\sigma}_{\mathbb{E}}^2 = \frac{1}{n_o - 1} \sum_{l=1}^{n_o} (\mathbb{E}[y_l] - \overline{\mathbb{E}[y_l]})^2 \quad (4.9)$$

$$\hat{\sigma}_{\sigma^2}^2 = \frac{1}{n_o - 1} \sum_{l=1}^{n_o} (\sigma_{y_l}^2 - \overline{\sigma_{y_l}^2}), \quad (4.10)$$

where  $\hat{\sigma}_{\mathbb{E}}^2$  and  $\hat{\sigma}_{\sigma^2}^2$  are the sample variances of  $\sigma_{\mathbb{E}}^2$  and  $\sigma_{\sigma^2}^2$  respectively, and  $\overline{\mathbb{E}[y_l]}$  and  $\overline{\sigma_{y_l}^2}$  are the sample means of the distributions of the expected values and the variances of the  $y_l$  respectively. The estimates,  $\hat{\sigma}_{\mathbb{E}}^2$  and  $\hat{\sigma}_{\sigma^2}^2$ , are then used in Equation 4.7 and Equation 4.8. As will be shown in Section 4.4,  $n_o$  is sufficiently large to neglect the uncertainty associated with these estimates. These intervals also require the estimation of operation-level expected values,  $\mathbb{E}[y_l]$ , and variances,  $\sigma_{y_l}^2$ , of  $\text{NO}_x$  emissions. These parameters, as will be discussed in Section 4.4, are estimated from a Monte Carlo simulation with a large number of model evaluations and thus, uncertainty associated with these estimates is also neglected.

As can be seen from Equation 4.7, as  $n_o$  approaches  $N_o$ , the confidence interval around  $\mathbb{E}[y_{tot}]$  narrows, eventually becoming a single point when  $n_o = N_o$ . Thus, there is a tradeoff between how many operations are analyzed in the surrogate model, and the tightness of the confidence intervals for the mean and variance of the total  $\text{NO}_x$ . Results from applying this method to construct confidence intervals for the mean and variance of the AEM  $\text{NO}_x$  output are presented in Section 4.4.

### 4.3.2 AEM surrogate for model development sensitivity analysis

The AEM has millions of factors: the six quantities shown in Table 4.1, defined for each flight segment of each operation. A sensitivity analysis could consider the sensitivities of each of these independently sampled factors individually, resulting in millions of sensitivity indices of limited use. From a practical standpoint, we are generally more interested in determining the sensitivity of model outputs to groups of factors. For example, for the  $\text{NO}_x$  emissions for an operation, we might wish to compute the contribution to variance of all fuelburn factors for that operation, where each factor is sampled independently across flight segments in the operation. Alternatively, for the total  $\text{NO}_x$  emissions summed over a set of operations, we might wish to compute the contribution to variance of all fuelburn factors for those operations, where again each factor is sampled independently across all flight segments. In this section, we present the extension of the global sensitivity analysis methodology to handle such cases. Further, application of the Central Limit Theorem, which is permissible given the additive nature of the AEM and ANOVA-HDMR, enables the calculation of confidence intervals around sensitivity indices computed using the AEM surrogate models in place of the full AEM. For the case of the AEM, distributional sensitivity analysis is not feasible given the need to group factors, which limits the ability to use acceptance/rejection sampling.

Consider the ANOVA-HDMR for the calculation of the  $\text{NO}_x$  emissions from a

single operation,  $l$ :

$$\begin{aligned}
y_l = & f_{0,l} + \sum_{k=1}^{N_s} f_{q_{kl}}(q_{kl}) + \sum_{k=1}^{N_s} f_{r_{kl}}(r_{kl}) + \sum_{k=1}^{N_s} f_{s_{kl}}(s_{kl}) + \sum_{k=1}^{N_s} f_{t_{kl}}(t_{kl}) \\
& + \sum_{k=1}^{N_s} f_{u_{kl}}(u_{kl}) + \sum_{k=1}^{N_s} f_{v_{kl}}(v_{kl}) + \text{interaction terms}, \tag{4.11}
\end{aligned}$$

where we use the input variable notation defined in Table 4.1. The term  $\sum_{k=1}^{N_s} f_{q_{kl}}(q_{kl})$  is the sum of all the single-factor functions of factor  $q_{kl}$ ; that is, the functions that depend only on the segment fuelburn inputs. The second summation is over those functions that depend only on the segment temperatures,  $r_{kl}$ , and so on for the other summations. Here, as in Equation 4.1,  $N_s$  segments have been assumed for operation  $l$ .

Since the goal is to compute sensitivities for inputs grouped across flight segments, we define  $q_l = \{q_{kl}\}_{k=1}^{N_s}$  to be the set of fuelburn segment inputs for operation  $l$ . Define  $r_l$ ,  $s_l$ ,  $t_l$ ,  $u_l$ , and  $v_l$  similarly for the other input quantities. Each summation in Equation 4.11 can then be written as

$$f_{q_l} = \sum_{k=1}^{N_s} f_{q_{kl}}(q_{kl}), \tag{4.12}$$

with analogous expressions defining  $f_{r_l}$ , etc. Then Equation 4.11 is written as

$$y_l = f_{0,l} + f_{q_l} + f_{r_l} + f_{s_l} + f_{t_l} + f_{u_l} + f_{v_l} + \text{interaction terms}. \tag{4.13}$$

Squaring and integrating Equation 4.13, as was done to arrive at Equation 3.6, gives

$$\text{var}(y_l) := D_{y_l} = D_{q_l} + D_{r_l} + D_{s_l} + D_{t_l} + D_{u_l} + D_{v_l} + \text{interaction partial variances}, \tag{4.14}$$

where  $D_{q_l}$  is the partial variance due to all fuelburn factors, and so on for the other factors.

Similarly, the AEM output  $y_{tot}$ , which, as noted in Section 4.1, is computed by aggregating the operational level outputs, is written in ANOVA-HDMR form by sum-

ming over the operations in Equation 4.11, which yields

$$\begin{aligned}
y_{tot} = & \sum_{l=1}^{N_o} f_{0,l} + \sum_{l=1}^{N_o} \sum_{k=1}^{N_s} f_{q_{kl}}(q_{kl}) + \sum_{l=1}^{N_o} \sum_{k=1}^{N_s} f_{r_{kl}}(r_{kl}) + \sum_{l=1}^{N_o} \sum_{k=1}^{N_s} f_{s_{kl}}(s_{kl}) + \sum_{l=1}^{N_o} \sum_{k=1}^{N_s} f_{t_{kl}}(t_{kl}) \\
& + \sum_{l=1}^{N_o} \sum_{k=1}^{N_s} f_{u_{kl}}(u_{kl}) + \sum_{l=1}^{N_o} \sum_{k=1}^{N_s} f_{v_{kl}}(v_{kl}) + \text{interaction terms.} \tag{4.15}
\end{aligned}$$

Now let  $q = \{q_l\}_{l=1}^{N_o}$  denote the set of fuelburn factors across all operations, and  $f_q = \sum_{l=1}^{N_o} \sum_{k=1}^{N_s} f_{q_{kl}}(q_{kl})$  be the sum of all the single-factor functions of all segment fuelburn factors, then Equation 4.15 is written as

$$y_{tot} = f_0 + f_q + f_r + f_s + f_t + f_u + f_v + \text{interaction terms,} \tag{4.16}$$

where  $f_0 = \sum_{l=1}^{N_o} f_{0,l}$  is the expected value of  $y_{tot}$  and the functions  $f_r$ ,  $f_s$ ,  $f_t$ ,  $f_u$ , and  $f_v$  are defined analogously to  $f_q$ . Squaring and integrating Equation 4.16 gives

$$\text{var}(y_{tot}) := D = D_q + D_r + D_s + D_t + D_u + D_v + \text{interaction partial variances,} \tag{4.17}$$

which may also be written as

$$D = \sum_{l=1}^{N_o} D_{q_l} + \sum_{l=1}^{N_o} D_{r_l} + \sum_{l=1}^{N_o} D_{s_l} + \sum_{l=1}^{N_o} D_{t_l} + \sum_{l=1}^{N_o} D_{u_l} + \sum_{l=1}^{N_o} D_{v_l} + \text{interaction partial variances.} \tag{4.18}$$

The total effect sensitivity index for  $q$ , denoted  $\tau_q$ , represents the relative contribution to the variance  $D$  of all fuelburn factors over all operations and segments. It can be computed using an analogous approach to that described in Section 3.2 as follows. As in Equation 3.10, we write

$$\tau_q = 1 - \frac{D_{q^c}}{D}, \tag{4.19}$$

where  $D_{q^c}$  is the sum of the variances due to all main effect terms and interaction effect terms that do not involve fuelburn. By breaking this expression into a sum over operations and using the fact that  $D = \sum_{l=1}^{N_o} D_{y_l}$ , the expression Equation 4.19 can



be written as

$$\tau_q = 1 - \frac{\sum_{l=1}^{N_o} (1 - \tau_{q_l}) D_{y_l}}{\sum_{l=1}^{N_o} D_{y_l}}, \quad (4.20)$$

where  $\tau_{q_l}$  is the total effect sensitivity index for  $q_l$ , the fuelburn inputs over operation  $l$ . Applying the Central Limit Theorem, it can be seen that Equation 4.20 is one minus a ratio of normal random variables. A similar derivation for main effect sensitivity indices leads to

$$S_q = \frac{\sum_{l=1}^{N_o} S_{q_l} D_{y_l}}{\sum_{l=1}^{N_o} D_{y_l}}. \quad (4.21)$$

As was the case for the expected values of  $\text{NO}_x$  emissions on the operational level in Section 4.3, the terms in Equation 4.20 and Equation 4.21 can be considered as samples from distributions. Therefore, to estimate the sensitivity indices given by Equation 4.20 and Equation 4.21, we apply the same process used to arrive at Equation 4.5 from Equation 4.3. In Equation 4.21 for example, the numerator,  $\sum_{l=1}^{N_o} S_{q_l} D_{y_l}$ , is equal to  $N_o \mathbb{E}[S_{q_l} D_{y_l}]$ , which may be estimated from  $\frac{N_o}{n_o} \sum_{l \in \mathcal{O}} [S_{q_l} D_{y_l}]$ . Just as in Equation 4.6, this estimate is normally distributed and converges to a single value when  $n_o = N_o$ . To estimate confidence intervals for  $\tau_q$  and  $S_q$ , we sample from distributions of the numerators and denominators to estimate the intervals empirically. It should be noted here that the confidence intervals computed for  $\tau_q$  and  $S_q$  will be conservative due to the fact that the numerator and denominator terms in both Equations 4.19 and 4.21 are positively correlated. By not including the correlation in the estimation of the confidence intervals, the estimate of the lower endpoint will be less than the true lower endpoint and the estimate of the upper endpoint will be greater than the true upper endpoint. This is due to the fact that the numerator in each equation must be less than or equal to the denominator in each equation, which leads to conservative intervals when the positive correlation of the terms is not included. Results from applying this method to the AEM sensitivity indices for the total emissions of  $\text{NO}_x$  are presented in the following section.

## 4.4 Results

A typical analysis run of the AEM consists of all operations conducted on a particular day that is considered a reasonable representative of all operations from a particular year. These days are referred to as *representative days*. The full AEM run for the representative day for the year 2005, which is the AEM model we consider here, has  $N_o = 68,343$  operations. Each of these operations requires a Monte Carlo simulation to calculate operation-level emissions outputs that are then aggregated, as shown in Equation 4.2 to produce the overall AEM output,  $y_{tot}$ . For the computational resources available for this study, a single model evaluation for one operation takes approximately  $2.31 \times 10^{-4}$  seconds. To perform both uncertainty and sensitivity analysis for a single operation requires 13 separate Monte Carlo simulations ( $2n + 1$  simulations, where the dimension of the factor space for the AEM is  $n = 6$ ), each of which consisted of 10,000 model evaluations in this study. Thus, running each operation of the AEM representative day for 2005 to perform uncertainty and sensitivity analysis on  $y_{tot}$ , would take approximately 570 hours, which is computationally expensive, especially if many different policy scenarios are to be considered. As will be shown in the following subsections, the methods presented in this chapter can be used to perform both uncertainty and sensitivity analyses on the AEM representative day with a surrogate model consisting of a randomly chosen subset of operations, while maintaining quantitative rigor in the analyses in a manner that is computationally efficient.

### 4.4.1 AEM surrogate results for decision-making uncertainty analysis

To estimate the confidence intervals for the mean and variance of the total  $\text{NO}_x$  emissions from the representative day, 9,914 operations were chosen randomly, without replacement, from the full set of operations.<sup>1</sup> As noted previously, a 10,000-iteration

---

<sup>1</sup>A total of 10,000 operations were chosen initially, however, 86 of these failed and did not produce meaningful results.

Table 4.2: 95% confidence intervals of the mean and variance of total NO<sub>x</sub> emissions computed with surrogate models of 2,500, 5,000, 7,500, and 9,914 operations.

$n_o$	Mean total NO <sub>x</sub> (gm)		Variance total NO <sub>x</sub> (gm <sup>2</sup> )	
	Lower	Upper	Lower	Upper
2500	$6.661 \times 10^9$	$7.926 \times 10^9$	$1.772 \times 10^{13}$	$2.925 \times 10^{13}$
5000	$6.888 \times 10^9$	$7.770 \times 10^9$	$1.963 \times 10^{13}$	$2.760 \times 10^{13}$
7500	$7.051 \times 10^9$	$7.782 \times 10^9$	$2.181 \times 10^{13}$	$2.885 \times 10^{13}$
9914	$6.899 \times 10^9$	$7.500 \times 10^9$	$2.073 \times 10^{13}$	$2.645 \times 10^{13}$

Monte Carlo simulation was run for each of the sampled operations, the results of which were used to compute operation-level NO<sub>x</sub> means and sample variances. These values were then used to estimate the expected value of the distribution of operation-level expected values of NO<sub>x</sub> emissions,  $\mu_E$ ; the variance of the expected value of the distribution of operation-level expected values of NO<sub>x</sub> emissions,  $\sigma_E^2$ ; the expected value of the distribution of operation-level variances of NO<sub>x</sub> emissions,  $\mu_{\sigma^2}$ ; and the variance of the distribution of operation-level variances of NO<sub>x</sub> emissions,  $\sigma_{\sigma^2}^2$ . As was noted in Section 4.3, these estimates are necessary for constructing the confidence intervals for the expected value and variance of the total NO<sub>x</sub> emissions of the full AEM and uncertainty in these estimates has been neglected. Figure 4-3 shows the behavior of these estimates as the number of operations in the subset,  $n_o$ , is increased from 2,500 to 9,914 operations.

The confidence intervals (95%) for the mean and variance of the total NO<sub>x</sub> emissions computed at values of  $n_o$  of 2,500, 5,000, 7,500, and 9,914, are presented in Table 4.2. Figure 4-4 presents the dependence of the confidence interval widths, in terms of percentage  $\pm$  of the surrogate model estimated values, for the mean and variance of total NO<sub>x</sub> emissions for a full run of the representative day as  $n_o$  increases from 2,500 to 9,914. These results show that by applying the surrogate modeling methodology described in Section 4.3 for uncertainty analysis in support of decision-making, confidence intervals for the mean and variance of total NO<sub>x</sub> emissions for the representative day can be constructed. These confidence intervals are quantitatively rigorous and display predictable convergence behavior that can be used to determine optimum tradeoffs between tighter intervals and longer run times.

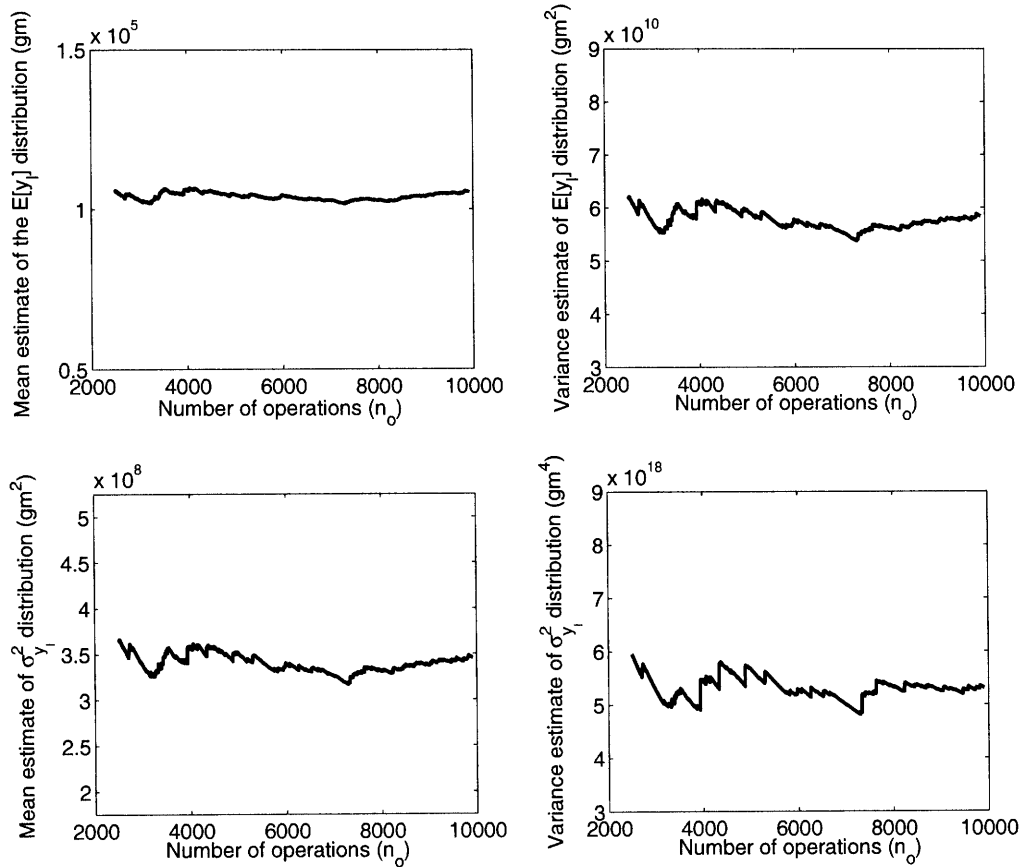


Figure 4-3: Estimates of  $\mu_{\mathbb{E}}, \sigma_{\mathbb{E}}^2, \mu_{\sigma^2},$  and  $\sigma_{\sigma^2}^2$  as the number of operations in the surrogate model,  $n_o$ , increases from 2,500 to 9,914.

#### 4.4.2 AEM surrogate results for model development sensitivity analysis

The total and main effect sensitivity indices were computed using the Sobol' method described in Section 3.2, applied to a surrogate model of  $n_o = 5,000$  operations sampled from the representative day. The resulting total and main effect sensitivity indices are shown in Figure 4-5. These results reveal that factors such as pressure and relative humidity can potentially be fixed for certain analyses since their total effect sensitivity indices are low, and that factors such as the reference emissions index of  $\text{NO}_x$  and temperature should be the focus of any future research aimed at trimming the variability in total  $\text{NO}_x$  emissions estimates from the AEM, since their main effect

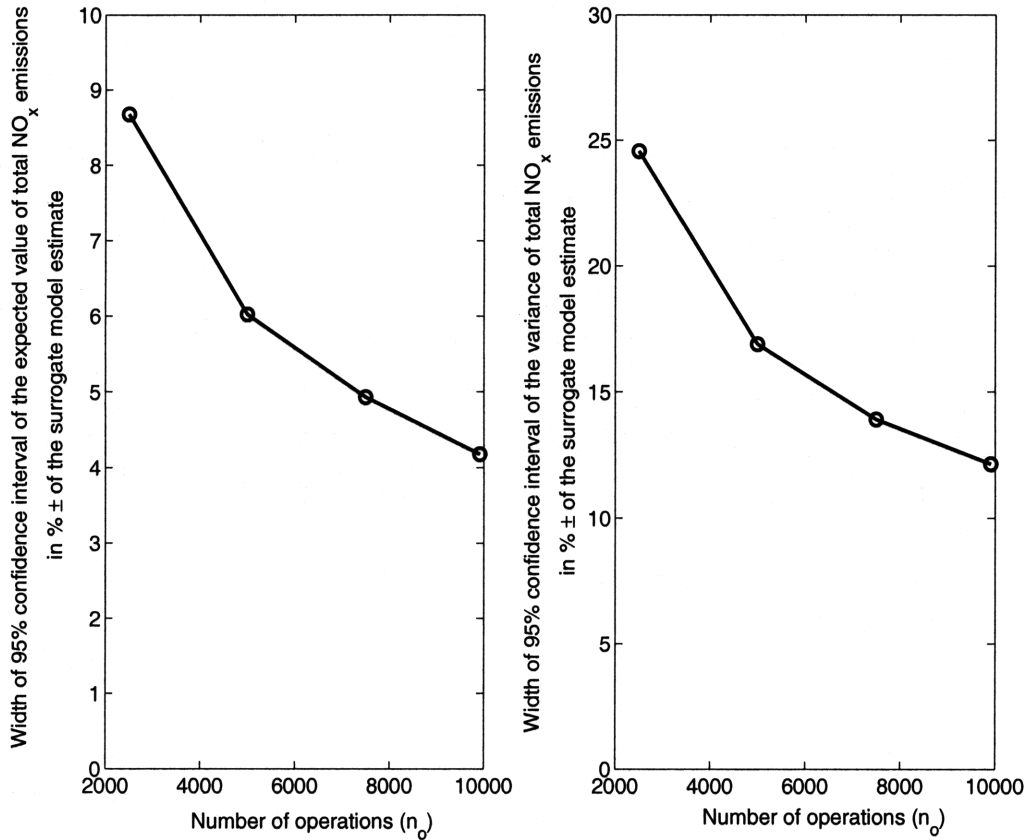


Figure 4-4: 95 percent confidence interval widths, in terms of percentage  $\pm$  of the estimated value, for the mean and variance of total  $NO_x$  emissions for a full run of the representative day as  $n_o$  increases from 2,500 to 9,914.

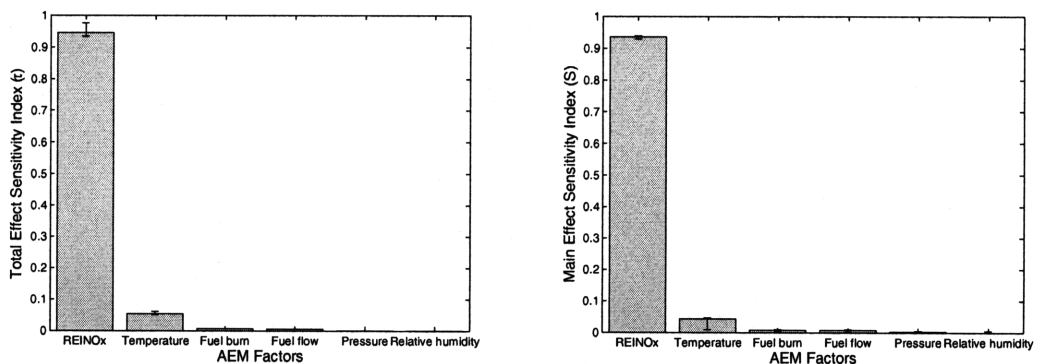


Figure 4-5: Total and main effect sensitivity indices for the AEM  $NO_x$  output as evaluated using a surrogate model consisting of 5,000 operations. The error bars show the 95% confidence intervals for each index.

sensitivity indices are highest.

These sensitivity results give valuable insight to guide model development; however, the question arises whether different conclusions might be drawn if the full model were used in place of the surrogate. In this situation, it is computationally impractical to use the full  $N_o = 68,343$  operations; however, the sensitivity results computed with the surrogate of  $n_o = 5,000$  operations can be rigorously bounded with confidence intervals using the methodology of Section 4.3. These confidence intervals were constructed by using a 10,000-iteration Monte Carlo simulation to compute each operation-level global sensitivity index required in Equation 4.20 and Equation 4.21. The intervals are shown for each sensitivity index in Figure 4-5.

Figure 4-6 shows the convergence behavior of the total effect sensitivity index of the temperature input. The convergence behavior of the other sensitivity indices is similar. Table 4.3 gives confidence intervals (95%) for the total and main effect

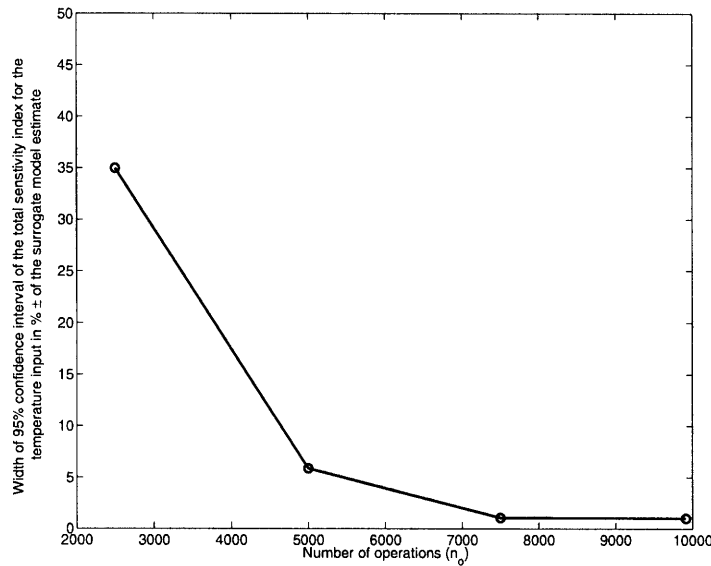


Figure 4-6: 95 percent confidence interval widths, in terms of percentage  $\pm$  of the surrogate model sensitivity index estimates, for the total and main effect sensitivity index of the temperature input as  $n_o$  increases from 2500 to 9914.

sensitivity indices for each input of the AEM for a full run of the representative day computed with a surrogate model of 5,000 operations.

Table 4.3: 95 percent confidence intervals of the total and main effect sensitivity indices for each factor of the AEM for a full run of the representative day computed with a surrogate model of 5,000 operations.

Input	Total effect sensitivity index		Main effect sensitivity index	
	Lower	Upper	Lower	Upper
REINO <sub>x</sub>	0.934	0.975	0.929	0.940
Temperature	0.050	0.053	0.009	0.060
Fuel burn	0.006	0.007	0.003	0.010
Fuel flow	0.005	0.006	0.001	0.010
Pressure	0.000	0.001	0.000	0.005
Relative Humidity	0.000	0.000	0.000	0.000

These results show that by applying the surrogate modeling methodology described in Section 4.3 for model development sensitivity analysis, confidence intervals for the global sensitivity indices of total NO<sub>x</sub> emissions for the full representative day can be constructed from a subset of operations. Just as for the confidence intervals constructed to support decision-making uncertainty analysis, these confidence intervals are quantitatively rigorous and display convergence behavior that can be used to determine optimum tradeoffs between tighter intervals and longer run times.





# Chapter 5

## Uncertainty Assessment of a Real-World Aviation Environmental Model

As stated in Chapter 1, an objective of this research is to *demonstrate* the application of the general approach methodology on a real-world model designed to support decision-making and policy-making processes. This chapter presents the application of the uncertainty assessment approach on the APMT-Impacts climate model, which is a component of the APMT-Impacts block shown in Figure 1-3. A detailed description of the model can be found in Refs. [3] and [66]. Section 5.1 provides background information on the model, which closely follows Refs. [3, 66]. Section 5.2 presents the uncertainty assessment of the model, and Section 5.3 discusses the general conclusions of this particular uncertainty assessment.

### 5.1 APMT-Impacts Climate Module

The APMT-Impacts climate *module* in the FAA environmental tools-suite discussed in Chapter 1, is responsible for estimating the impacts of aviation on global climate. The term *module* is used here to note the fact that many different climate models may be used in the tools-suite, such as conservative, nominal, and low impact models

[67]. These different models, as well as motivation and background on the modeling methodology, will be discussed in more detail in the following subsections.

### **5.1.1 Motivation for the APMT-Impacts climate module**

The FAA tools-suite, discussed in Chapter 1, exists in part to answer questions regarding how to evaluate interdependencies among local air quality, noise, and climate impacts due to aviation. Policy-makers, in an effort to balance society’s economic and environmental needs, wish to assess the full impact of candidate policies, while accounting for potential interdependencies. This results in a complex decision-making setting, in which policy-makers need a “shared conception of what is at stake in the choice of one level of effort or another, and a common terminology for incorporating these considerations into international negotiations and domestic decision-making” [68]. For this reason, the FAA-tools suite presents benefits and costs of policies in monetary terms to provide a common basis for evaluating interdependencies. The role of the APMT-Impacts climate module in the tools-suite is thus the estimation of the marginal climate impacts of new aviation activities and the valuation of these impacts in monetary terms.

### **5.1.2 Modeling methodology of the climate module**

The modeling approach of the APMT-Impacts climate module is shown in Figure 5-1, which follows the approaches of Refs. [69, 70, 71, 72, 73]. The method begins with the estimation of current and future emissions inventories for both aviation and all anthropogenic sources. The potential change in globally-averaged surface temperature is then estimated from impulse response functions, which conceptualize a year of emissions as an impulse, and are derived from carbon-cycle and general circulation models [74, 75, 76]. From the change in globally-averaged surface temperature, metrics which can be used in decision-making, such as damage as a percent impact on gross domestic product (GDP), and the net present value (NPV) of climate change due to aviation, are estimated. The impacts of emissions over the entire period dur-

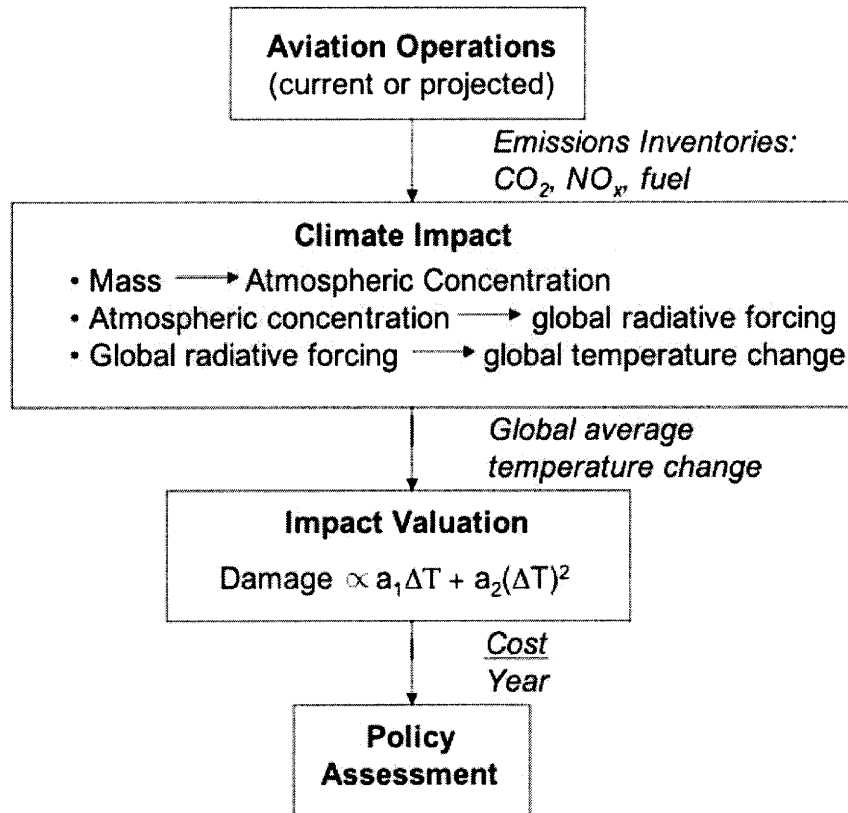


Figure 5-1: Overview of the modeling approach for the APMT-Impacts climate module (adapted from [3]).

ing which significant effects persist are considered. For the case of CO<sub>2</sub>, this period is several centuries. What follows is a more detailed description of each step of the modeling methodology shown in Figure 5-1.

### Aviation Operations

Current estimates of aviation operations can be input to the climate module in many ways. Within the FAA tools-suite, they are typically estimated as a one year aviation emissions impulse based on 2003 data from AEDT, followed by a 30 year projection of aviation emissions. Background CO<sub>2</sub> and GDP growth scenarios can also be estimated in many ways, and are usually based on the International Panel on Climate Change (IPCC) Special Report on Emissions Scenarios (SRES) scenario A1B [77]. The scenario is used out to 2100, after which the emissions are assumed to remain constant and the GDP is extrapolated linearly based on the last decade of the scenario.

## Climate Impact

The climate module estimates the impacts of aviation on the climate through both CO<sub>2</sub> and non-CO<sub>2</sub> related effects. The impact of CO<sub>2</sub> on the atmosphere is non-linear, meaning additional units of CO<sub>2</sub> cause progressively less radiative forcing, where radiative forcing (RF) is a perturbation in the energy balance system of the Earth by incoming and outgoing radiation in the atmosphere. Thus, the impact of aviation related CO<sub>2</sub> is determined by calculating the impact of all anthropogenic sources and subtracting the impact of all anthropogenic sources with aviation removed:

$$\begin{aligned} \text{Impact}(\text{CO}_2 \text{ aviation}) = & \quad \text{Impact}(\text{CO}_2 \text{ anthropogenic}) - \\ & \quad \text{Impact}(\text{CO}_2 \text{ anthropogenic} - \text{aviation}). \end{aligned} \quad (5.1)$$

The change in concentration of atmospheric CO<sub>2</sub> due to anthropogenic sources is estimated from impulse response functions derived from carbon-cycle models following the approach of Ref. [70]. The calculation proceeds as follows:

$$\begin{aligned} G_c(t') &= \sum_{j=1}^{n_j} \alpha_j e^{-(t')/\tau_j} \\ \Delta X_{\text{CO}_2}(t') &= \int_{t_0}^{t'} Q_{\text{CO}_2}(t'') \cdot G_c(t' - t'') dt'' \\ &\approx \sum_{n=0}^{N-1} Q_{\text{CO}_2}(t_0 + n\Delta t) \cdot G_c(t' - t_0 - n\Delta t) \cdot \Delta t \\ N &= (t' - t_0)/\Delta t, \end{aligned} \quad (5.2)$$

where  $G_c(t')$  is the carbon cycle impulse response function,  $Q_{\text{CO}_2}(t'')$  is the mass of CO<sub>2</sub> emitted from anthropogenic sources,  $\Delta X(t')$  is the corresponding change in atmospheric CO<sub>2</sub> concentration, and  $\Delta t$  is typically defined as one year. The terms  $\alpha_j$ ,  $n_j$ , and  $\tau_j$  are specific to the carbon cycle impulse response function employed, which will be discussed further in the following subsection.

The resulting normalized radiative forcing,  $RF_{\text{CO}_2}^*$  at time  $t'$ , is associated with

CO<sub>2</sub> concentration at time  $t'$  by assuming a logarithmic dependence [70]:

$$RF_{CO_2}^* = \log_2 \left( \frac{X_{CO_2(\text{present})} + \Delta X_{CO_2}(t')}{X_{CO_2(1750)}} \right), \quad (5.3)$$

where the present atmospheric concentration of CO<sub>2</sub> is  $X_{CO_2(\text{present})}$  and the atmospheric concentration of CO<sub>2</sub> in the year 1750 is  $X_{CO_2(1750)}$ , which is taken as 278 ppmv [78].

The global average temperature change due to CO<sub>2</sub>,  $\Delta T_{CO_2}(t)$ , from some time  $t_0$  to time  $t$ , is then determined from a simple energy balance model as follows:

$$\begin{aligned} \Delta T(t) &= \frac{1}{C} \int_{t_0}^t \Delta F(t') \exp\left(\frac{t' - t}{C\lambda^*}\right) dt' \\ \lambda^* &= \frac{\lambda}{RF_{2XCO_2}} \\ \Delta F(t') &= RF^* \cdot RF_{2XCO_2} \\ \tau &= C\lambda^*, \end{aligned} \quad (5.4)$$

where  $C$  ( $4.2 \times 10^8$  J/Km<sup>2</sup>) is the ocean heat capacity for a global ocean mixed layer of 100m depth,  $RF_{2XCO_2}$  ( $3.7$  W/m<sup>2</sup>) is the radiative forcing for a doubling of CO<sub>2</sub> relative to pre-industrial levels [78],  $\lambda$  is the climate sensitivity,  $RF^*$  is the normalized radiative forcing for the different effects,  $\Delta F$  is the radiative forcing due to the effects of different aviation emissions, and  $\tau$  is the time constant of the climate system.

The non-CO<sub>2</sub> related effects of aviation emissions consist of both short-lived and longer time-scale impacts. For the short-lived effects (i.e. effects assumed to have lifetimes of one year or less, including the effect of NO<sub>x</sub> on ozone, contrails and aviation-induced cirrus, water, sulfates, and soot) it is assumed that the radiative forcing is active only in the year of the emissions. The temperature change associated with each effect is determined in the same manner as for CO<sub>2</sub>. Following Ref. [70], aviation short-lived effects are represented by scaling the normalized radiative forcing

for each effect relative to CO<sub>2</sub>:

$$RF_{short_j}^*(t_0) = \frac{\lambda_{short_j}}{\lambda_{CO_2}} \cdot \frac{RF_{short_j}^{ref}}{RF_{2XCO_2}} \cdot \frac{Q_{short_j}(t_0)}{Q_{short_j}^{ref}}, \quad (5.5)$$

where the  $\lambda$ 's are the sensitivities for each effect, the  $RF^{ref}$ s are the reference radiative forcing values for each effect,  $RF_{2XCO_2} = 3.7 \text{ W/m}^2$  is the radiative forcing for a doubling of atmospheric CO<sub>2</sub> concentration relative to the pre-industrial level,  $Q_{short}$  is the emissions quantity for each effect, and  $Q_{short}^{ref}$  is the reference emissions quantity corresponding to  $RF_{short}^{ref}$ . The ratio,  $\lambda_{short}/\lambda_{CO_2}$  is the efficacy for a given effect.

The longer time-scale impacts considered are the decrease in methane lifetime resulting from the presence of NO<sub>x</sub>, and the resulting small decrease in ozone over the same time period. An initial value for the methane radiative forcing is derived by scaling the values for short-lived ozone from Ref. [70] using 100-year time integrated radiative forcing results from Ref. [79], as shown in Equation 5.6.

$$\begin{aligned} \frac{\int_0^{100} RF_{ozone-short}(t) dt}{\int_0^{100} RF_{methane}^{ref}(t_0) \exp(-t/\tau_{methane}) dt} &= \frac{RFyr_{ozone-short}}{RFyr_{methane}} \\ \int_0^{100} RF_{ozone-short}(t) dt &\approx RF_{ozone}^{ref} \cdot 1yr \\ RF_{methane}^{ref} &= \frac{RFyr_{methane}}{RFyr_{ozone-short}} \cdot \frac{RF_{ozone}^{ref} \cdot 1yr}{\tau_{methane}}. \end{aligned} \quad (5.6)$$

The mean 100-year  $RFyr$  values are taken to be 5.06 mW-yr/m<sup>2</sup> for ozone and -4.00 mW-yr/m<sup>2</sup> for methane. The mean  $e$ -folding time for methane,  $\tau_{methane}$ , which is the time in which methane increases by a factor of  $e$ , is set at 11.07 yrs as defined by Ref. [79]. A similar procedure is used to estimate the radiative forcing for the longer-term NO<sub>x</sub>-ozone cooling impact.

## Impact Valuation

The climate module uses the damage function defined in Ref. [73] to relate temperature change to economic cost. The damage function calculates a financial loss as a percentage of GDP from a weighted sum of surface temperature change and squared

surface temperature change as follows:

$$D(t) = a_1 \Delta T_{1900}(t) + a_2 \Delta T_{1900}(t)^2, \quad (5.7)$$

where  $\Delta T_{1900}$  is the temperature change since 1900,  $a_1$  and  $a_2$  are coefficients, and  $D(t)$  is the estimated damage in percentage of GDP. A reference temperature change from 1900 to 2003 is also used in the estimates to allow for damage from aviation from 2003 onwards to be estimated.

The final step is to discount the calculated damage costs and determine the NPV of the climate impact. This is done using Equation 5.8

$$\text{NPV}(n) = \sum_{n'=n_0}^n \frac{d(n' - n_0)}{(1 + r)^{n' - n_0}}, \quad (5.8)$$

where  $d(n' - n_0)$  is the estimated damage costs during the time interval  $(n' - n_0)$ ,  $r$  is the discount rate and  $n$  is the period over which the discount rate is applied.

### 5.1.3 The nominal climate model

As noted previously, the climate *module* in the FAA environmental tools-suite, is responsible for estimating the impacts of aviation on global climate. The module is exercised by considering different climate models for different sets of assumptions. These assumptions are such things as how the discount rate is set, what is used as a carbon-cycle model, what is used as a temperature response model, and how the damage coefficients are set. As mentioned previously, in some studies (e.g. [33]), these assumptions and parameters are deemed to have epistemic modeling uncertainty. In these studies, the assessment process consists of a double-loop approach, where modeling epistemic uncertainties are sampled from in an outer-loop, and aleatory uncertainties are sampled from in an inner loop, which results in families of cumulative distribution functions. The work being done on the FAA tools-suite however, takes a different approach to dealing with epistemic modeling uncertainties. As mentioned in Section 2.1.4, epistemic modeling uncertainties are treated by considering sets of

combinations of interest, which are referred to as *lenses* [67]. Each lens represents a different method of modeling the impacts of aviation on climate. Some examples are the conservative model, where all factors are set to values that cause the largest impacts of aviation on climate change, the nominal model, where factors are treated as random variables, and the low impact model, where all factors are set to values that cause the smallest impacts of aviation on climate change. Figure 5-2 presents a tornado chart representation of the impacts of different realizations of modeling epistemic uncertainty on the climate module output of NPV. The impacts are presented by shifting, on a one-at-a-time basis, each factor uncertainty (e.g. damage coefficients and climate sensitivity) and each modeling uncertainty (e.g. discount rate, which carbon cycle model to employ, etc.) from the nominal value to a minimum and maximum value for the factor uncertainties, and to different possible realizations for the modeling uncertainties. The uncertainty assessment approach established in this work,

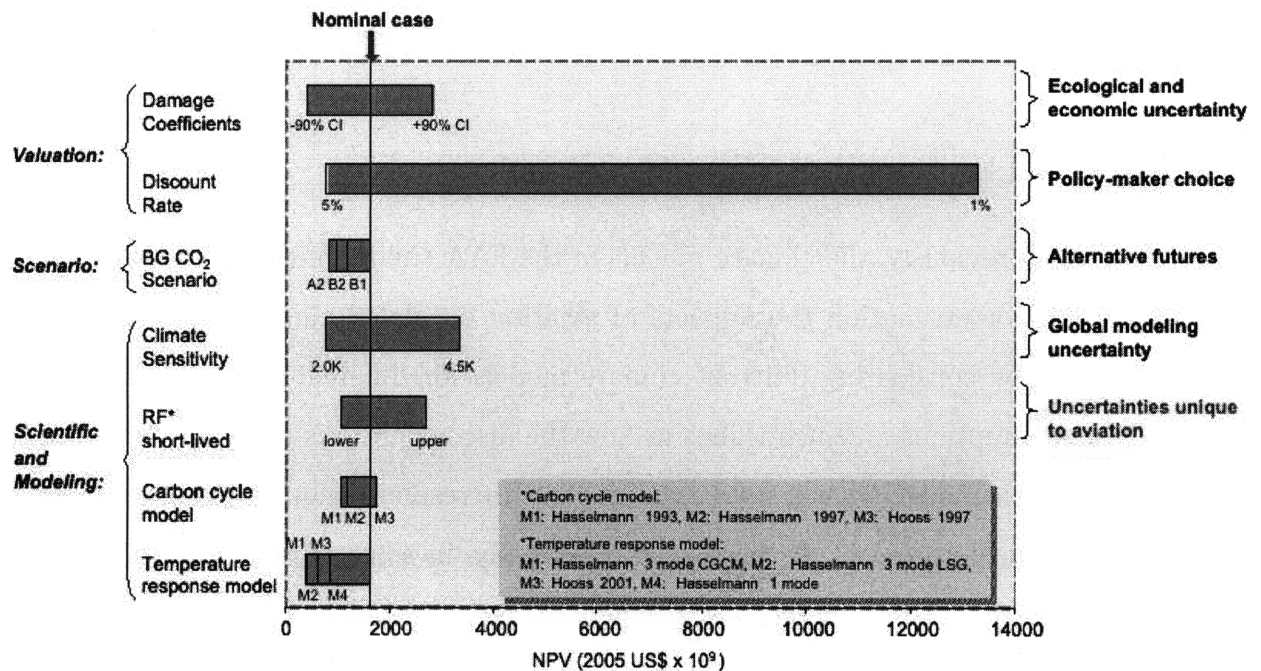


Figure 5-2: Tornado chart of the NPV output of the climate module, which shows the impacts of epistemic modeling uncertainty.

as noted previously, does not deal with epistemic modeling uncertainty, and instead focuses only on the uncertainties present within a single lens. The lens for which the



uncertainty assessment presented in this chapter focuses on is the referred to as the nominal climate model.

The nominal climate model can analyze the impacts of aviation on climate change for many different aviation scenarios, such as different policy options and a baseline scenario, which is the expected scenario if no policy actions are taken. The assessment presented here focuses entirely on the baseline scenario. It is the topic of future work to consider various policy scenarios and the differences between baseline and policy scenarios.

## **5.2 Uncertainty Assessment of the APMT-Impacts Nominal Climate Model**

The APMT-Impacts nominal climate model, as mentioned earlier, is used to estimate the impacts of aviation on the environment in a manner that suits the needs of the policy-making community. Given that the model is exercised in a policy-making context, the ability to inform analyses with uncertainty associated with model outputs and the identification of major contributors to output variability are considered key priorities [1]. An uncertainty assessment with these priorities in mind, is presented for the nominal climate model in the following subsections. The assessment follows the steps detailed in Chapter 2. Step 7, the presentation of results, is not discussed in its own subsection, since results are presented for uncertainty analysis and sensitivity analysis in their respective subsections.

### **5.2.1 Step 1: Establishing uncertainty assessment objectives**

The objectives of an uncertainty assessment for complex models intended to support decision- and policy-making processes, as established previously in Section 2.1.1, are as follows:

- Goals for supporting decision-making processes

1. Provide quantitative evaluation of the performance of the model relative to fidelity requirements for various analysis scenarios.
  2. Provide quantitative comparisons of various policy scenarios, taking into account uncertainty in model outputs.
- Goals for furthering the development of the model
    1. Identify gaps in functionality that significantly impact the achievement of model requirements, leading to the identification of high-priority areas for further development.
    2. Rank factors based on contributions to output variability to inform future research and validation efforts.

These are the objectives of the uncertainty assessment of the nominal climate model.

## 5.2.2 Step 2: Document assumptions and limitations of the model

As discussed in Chapter 2, the assumptions and limitations of the model should be documented, with such information as what each assumption is and what it means, what the assumption affects, what does the assumption imply in terms of model validity, and what are the implications of the assumption in terms of model applicability. The assumptions and limitations of the nominal climate model are documented here. The information provided here is used, where possible, to inform the uncertainty characterization of certain factors in Step 3, and also to make it clear how model results can be applied.

### Nominal Climate Model Assumptions and Limitations

1. *Use of impulse response functions*
  - **Assumption:** Aviation related climate impacts can be estimated using a two-step process involving impulse response functions. First, aviation CO<sub>2</sub> emissions are translated to changes in atmospheric concentrations

of CO<sub>2</sub> using a carbon-cycle impulse response model to estimate radiative forcing. Next, changes in radiative forcing are related to changes in globally-averaged temperature using a simplified energy balance model.

- **What it means:** This approach ignores feedbacks between the carbon-cycle and general circulation model (GCM) impulse response functions, though given the small perturbation in CO<sub>2</sub> emissions due to aviation, these feedbacks are not expected to be substantial. Ideally, estimates of aviation related impacts would be derived using complex three-dimensional general circulation models that take into account different feedback mechanisms involving the flow of carbon between the ocean, atmosphere, and terrestrial components, including the biosphere. Rising global temperatures would alter the behavior of the different components of the carbon-cycle which would in turn provide a positive or negative feedback to the trends in temperature rise.
- **What it affects:** This assumption affects the estimated temperature change, which is used in the estimation of both damage and NPV.
- **Assessing the impacts on model validity:** It is difficult to assess the validity of the two-step process used in the nominal climate model since complex GCMs that account for different feedbacks and produce higher fidelity results tend to be computationally expensive and thus, difficult to compare with. Additionally, the lack of scientific understanding on aviation impacts such as contrails and induced cirrus makes it difficult to provide factors to the GCMs to obtain impact estimates.
- **Implications for model applicability:** It is well-known that the two-step process involving impulse response functions and simple energy balance models may not capture all the carbon-cycle feedbacks, thus, for policies where feedback mechanisms are expected to play an important role, this assumption will limit the applicability of the model.

## 2. *Bern carbon-cycle model*

- **Assumption:** Complex three-dimensional global carbon-cycle models are represented by a simplified impulse response function known as the Bern carbon-cycle model. This model has been calibrated against existing sophisticated carbon-cycle models and was used in the fourth assessment report of the IPCC [80].
- **What it means:** This assumption dictates the values used for the coefficients of the carbon-cycle model given by Equation 5.2.
- **What it affects:** The Bern carbon-cycle model determines the change in atmospheric CO<sub>2</sub> concentration due to all anthropogenic emissions, including aviation. The estimated CO<sub>2</sub> concentration is then used to estimate the globally-averaged temperature change, which is then used to estimate damage and NPV.
- **Assessing the impacts on model validity:** As discussed for the previous assumption, it is not possible to fully-understand the impacts of using the Bern carbon-cycle model to estimate atmospheric CO<sub>2</sub> concentrations without further research.
- **Implications for model applicability:** The use of the Bern carbon-cycle model in place of complex high fidelity carbon-cycle models implies that the nominal climate model cannot take in to account all carbon-cycle feedback mechanisms. For policies where feedback mechanisms are expected to play an important role in the estimation of atmospheric CO<sub>2</sub> concentrations, this model assumption will limit the applicability of the nominal climate model.

## 3. *Simple energy balance temperature response model*

- **Assumption:** The nominal climate model assumes that temperature change resulting from radiative forcing changes can be estimated from a simple energy balance model, as given by Equation 5.4.

- **What it means:** This assumption implies that climate sensitivity, which is the global average air temperature change that would result from a sustained doubling of atmospheric CO<sub>2</sub> concentration, is the key driver in the estimation of temperature change from radiative forcing changes.
- **What it affects:** The energy balance temperature response model affects the estimated globally-averaged temperature change, which is then used to estimate both damage and NPV.
- **Assessing the impacts on model validity:** As with the use of carbon-cycle impulse response models, the impact of using a simplified energy balance model cannot be fully understood without further research.
- **Implications for model applicability:** The simplified energy balance model was derived for climate projections over a long time frame. The application of the model to a single years worth of emissions is thus unvalidated, and results and inferences made using model outputs should include a disclaimer related to this assumption.

#### 4. *Short-lived non-CO<sub>2</sub> effects*

- **Assumption:** Radiative forcings of non-CO<sub>2</sub> emissions, with the exception of long-lived NO<sub>x</sub> impacts on methane and the secondary effect of methane on ozone, are only active in the year of the emissions, and it is assumed that short-lived emissions impacts are independent of each other and independent of background emissions.
- **What it means:** Following Ref. [70], aviation short-lived impacts are represented by scaling the normalized RF for different climate responses relative to CO<sub>2</sub>, and these short-lived effects have lifetimes on the order of one year.
- **What it affects:** This assumption affects the normalized RFs, which affect the temperature change due to short-lived emissions, which affects the overall estimate of temperature change, and thus the estimates of damage

and NPV.

- **Assessing the impacts on model validity:** There has not been a study aimed at determining how this assumption impacts model results.
- **Implications for model applicability:** The assumption is standard in the literature, however, if the short-lived effects are found to last for substantially longer or shorter time periods, the assumption will need to be revisited.

#### 5. *Use of Globally-Averaged Metrics*

- **Assumption:** Globally-averaged metrics are used to represent the physical impacts of aviation on climate.
- **What it means:** The model does not have the capability to estimate regional climate impacts due to aviation activity.
- **What it affects:** It is a core assumption of the model and affects all results and inferences that can be made with the model.
- **Assessing the impacts on model validity:** At present, there are no impulse response functions available in the literature that allow for estimating regional impacts. Further research is necessary to assess the assumption that impacts can be estimated reasonably well through global averaging.
- **Implications for model applicability:** It is well-understood that RF due to contrails, aviation-induced cirrus cloudiness, and production of ozone via  $\text{NO}_x$  will occur in regions where aircraft fly, which is predominantly the northern hemisphere. Modeling such RF as globally-uniform, and assuming that it may be simply superposed with RF due to well-mixed gases such as  $\text{CO}_2$  and methane, may inaccurately represent the more complex response of the climate to spatially non-homogeneous forcing. However, globally-averaged impact estimates remain the most widely used measure for communicating climate change effects (e.g. in IPCC reports).

## 6. *Use of scenario A1B to project future anthropogenic activity*

- **Assumption:** To estimate marginal climate impacts of future aviation activity, the full time horizon over which the impacts last should be taken into account. Since aviation-related CO<sub>2</sub> concentrations changes are calculated relative to background atmospheric CO<sub>2</sub> levels, future projections of anthropogenic activities are necessary to determine trends in atmospheric CO<sub>2</sub> levels. Additionally, future projections of global GDP are also necessary to estimate monetary impacts. These future projections are based on the IPCC SRES scenario A1B.
- **What it means:** Estimates of aviation CO<sub>2</sub> related temperature change and resulting damages to the global GDP are dependent on the assumed future scenario.
- **What it affects:** The assumption affects the estimated CO<sub>2</sub> related temperature change, which affects estimates of damage and NPV.
- **Assessing the impacts on model validity:** It is not possible to know what will happen in the future, thus a future scenario must be selected. Model results are then considered valid under an assumed future.
- **Implications for model applicability:** This assumption is one of the key drivers for the development of lenses for the climate module. The nominal climate model, which uses scenario A1B, is applicable to those policy-makers who believe scenario A1B is a reasonable representation of what will happen in the future. For policy-makers who do not believe scenario A1B is reasonable, an alternate lens should be used.

## 7. *Use of the damage function*

- **Assumption:** The damage function of Ref. [73] is applied to relate the global mean surface temperature change to welfare loss as a fraction of global GDP.

- **What it means:** The nominal climate model assumes the damage to global GDP caused by aviation can be completely characterized by the change in globally-averaged temperature and the relation given by Equation 5.7.
- **What it affects:** The assumption affects the estimates of damage and NPV.
- **Assessing the impacts on model validity:** There has been a complete study of this assumption and it is recognized generally as the most extensive and detailed to date [81]. However, other functional forms are being studied, as well as wider ranges of the current coefficients.
- **Implications for model applicability:** The approach taken by the nominal climate model for damage valuation has been criticized for the simplifying assumptions it contains, such as excluding non-market impacts (e.g. loss of natural beauty and extinction of species). Users of model results must understand that these impacts are not taken into account when looking at the final NPV output.

## 8. *Discount Rate*

- **Assumption:** The nominal climate model evaluates the NPV of climate impacts by assuming a constant discount rate of 0.035.
- **What it means:** Discounting will cause effects in the future to be valued less than effects in the present time. The use of a constant discount rate may lead to underestimated distant future impacts. A constant discount rate also assumes no relation to GDP growth scenarios.
- **What it affects:** The assumption affects the NPV estimates.
- **Assessing the impacts on model validity:** This step in the modeling is considered valid under the assumed discount rate. If another method of discounting is preferred, or another value of the discount rate is more appropriate, a different climate module lens should be used.



- **Implications for model applicability:** As stated above, the model is applicable to those policy-makers who believe the discount rate should be set to 0.035. For those that do not believe the discount rate should be set to 0.035, a different climate module lens should be used.

9. *Impacts are estimated 800 years into the future*

- **Assumption:** The effects of aviation on climate are negligible after 800 years.
- **What it means:** Estimates of global temperature change and damage are calculated for 800 years into the future, which are then used to estimate the NPV of aviation impacts on climate.
- **What it affects:** This assumption affects the NPV estimate of the nominal climate model.
- **Assessing the impacts on model validity:** Ref. [66] presents results that show that the impacts of aviation on temperature change are negligible after 800 years.
- **Implications for model applicability:** For the emissions currently considered in the nominal climate model, there are no restrictions on model applicability, however, if other emissions are considered, or if the estimation procedures for the impacts of the current emissions are modified, this assumption will have to be revisited.

As noted previously, the information provided here is essential to the proper application of the nominal climate model. The rest of the uncertainty assessment of the model, which is discussed in the following subsections, quantifies the uncertainty in model outputs due to the uncertainties associated with the factors of the climate model. Thus, it must be understood that the quantitative uncertainty assessment results quantify the uncertainty associated with model factors with each of the above assumptions in place, and therefore, do not quantify the uncertainty associated with the assumptions themselves.

### 5.2.3 Step 3: Document factors and outputs

As discussed in Chapter 2, the factors of the model should be documented, along with information regarding factor units, which outputs the factor affects, and any known uncertainty information regarding the factor. Outputs should also be documented, including units, possible downstream use, and upon which factors the output depends. As noted in Chapter 1, an *output* is defined as a model result of interest. For this particular uncertainty assessment of the nominal climate model, the focus will be on the NPV estimate only for brevity. Table 5.1 presents the factor information required for an uncertainty assessment and Table 5.2 presents the output information.

### 5.2.4 Step 4: Classify and characterize uncertainty

Based on the uncertainty information and the sources of that information, the type of uncertainty associated with each factor, and the distribution associated with each factor is given in Table 5.3. Uniform distributions are represented as  $U[a, b]$ , where  $a$  is the minimum and  $b$  is the maximum value. Triangular distributions are represented as  $T(a, b, c)$ , where  $a$  is the minimum value,  $b$  is the maximum value, and  $c$  is the mode of the distribution. Normal distributions are represented as  $\mathcal{N}(\alpha, \beta)$ , where  $\alpha$  is the mean, and  $\beta$  is the variance.

Each of the factors of the nominal climate model has been classified here as having epistemic uncertainty associated with it since, while each factor is expected to contain some amount of aleatory uncertainty, there is currently limited knowledge regarding those uncertainties.

### 5.2.5 Step 5: Conduct uncertainty analysis

An uncertainty analysis was conducted on the nominal climate model by propagating factor uncertainty through the model using Monte Carlo simulation as discussed in Chapter 2. The mean and standard deviation of the NPV estimate are given in Table 5.4. Figure 5-3 presents a histogram of the NPV estimates that resulted from the Monte Carlo simulation. This information can be used in a variety of ways;

Table 5.1: Nominal climate model factors and uncertainty information.

Factor	Units	Uncertainty Information	Source	Outputs
AEDT 2005 CO <sub>2</sub> Emissions	g	+/- 5%	[3]	NPV
AEDT 2005 Fuelburn	g	+/- 5%	[3]	NPV
AEDT 2005 NO <sub>x</sub> Emissions	g	+/- 10%	[3]	NPV
Climate Sensitivity	K	range: [2.0, 4.5] most likely: 3.0	[80]	NPV
Ref. temperature change	K	range: [0.4, 0.8] most likely: 0.6	[78]	NPV
RF doubling of CO <sub>2</sub>	W/m <sup>2</sup>	range: [3.5, 4.1] most likely: 3.7	[78]	NPV
RF NO <sub>x</sub> -O <sub>3</sub>	mW/m <sup>2</sup>	range: [0,35] most likely: 21.9	[70]	NPV
RF H <sub>2</sub> O	mW/m <sup>2</sup>	range: [0,6] most likely: 2	[70]	NPV
RF sulfate	mW/m <sup>2</sup>	range: [-7.5,0] most likely: -3.5	[70]	NPV
RF soot	mW/m <sup>2</sup>	range: [0,5] most likely: 2.5	[70]	NPV
RF contrails	mW/m <sup>2</sup>	range: [0,30] most likely: 10	[70]	NPV
RF cirrus	mW/m <sup>2</sup>	range: [0,80] most likely: 30	[70]	NPV
NO <sub>x</sub> -O <sub>3</sub> Efficacy	unitless	range: [0.75,1]	[82]	NPV
H <sub>2</sub> O Efficacy	unitless	value: 1	[82]	NPV
Sulfate Efficacy	unitless	range: [0.68,1.09]	[82]	NPV
Soot Efficacy	unitless	range: [0.62,1.29]	[82]	NPV
Contrails Efficacy	unitless	range: [0.59,1]	[82]	NPV
Cirrus Efficacy	unitless	value: 1	[82]	NPV
Damage Coefficients	%GDP/K %GDP/K <sup>2</sup>	value $a_1 = 0$ mean $a_2 = 0.0028388$ std. dev. $a_2 = 0.0013$	[73]	NPV

Table 5.2: Nominal climate model output information for uncertainty assessment.

Output	Units	Factors
NPV	US\$B2005	All factors

depending on how the information is used, the variability in the NPV estimate may cause the results to be of limited use. For example, the results could be used to for comparison with another output, such as with the impacts of aviation on air quality, or

Table 5.3: Uncertainty classification and characterization of nominal climate model factors.

Factor	Uncertainty Type	Distribution
AEDT CO <sub>2</sub> Emissions	Epistemic	$U[-5\%, +5\%]$
AEDT Fuelburn	Epistemic	$U[-5\%, +5\%]$
AEDT NO <sub>x</sub> Emissions	Epistemic	$U[-10\%, +10\%]$
Climate Sensitivity	Epistemic	$T(2.0, 4.5, 3.0)$
Ref. Temperature Change	Epistemic	$T(0.4, 0.8, 0.6)$
RF doubling of CO <sub>2</sub>	Epistemic	$T(3.5, 4.1, 4.7)$
RF NO <sub>x</sub> -O <sub>3</sub>	Epistemic	$T(0, 35, 21.9)$
RF H <sub>2</sub> O	Epistemic	$T(0, 6, 2)$
RF Sulfate	Epistemic	$T(-7.5, 0, -3.5)$
RF Soot	Epistemic	$T(0, 5, 2.5)$
RF Contrails	Epistemic	$T(0, 30, 10)$
RF Cirrus	Epistemic	$T(0, 80, 30)$
NO <sub>x</sub> -O <sub>3</sub> Efficacy	Epistemic	$U[0.75, 1]$
Sulfate Efficacy	Epistemic	$U[0.68, 1.09]$
Soot Efficacy	Epistemic	$U[0.62, 1.29]$
Contrails Efficacy	Epistemic	$U[0.59, 1]$
Damage Coefficient $a_2$	Epistemic	$\mathcal{N}(0.0028388, 0.0013^2)$

Table 5.4: Estimates of the mean and standard deviation of the Nominal climate model NPV for a years worth of emissions from aviation.

	mean	standard deviation
NPV Billion US\$2005	2038.30	1067.60

noise related effects. If the impacts on climate are substantially higher or substantially lower than the impacts on air quality, where substantially implies the histograms of the two outputs do not overlap significantly, then it may not be necessary to perform a sensitivity analysis to determine which factors should be researched to help reduce the variability. However, if the case were that the distributions of the impacts on climate and on air quality had substantial overlap, it may be necessary to refine the estimates of each to make a better comparison. Here, the standard deviation is about 50% of the mean of the NPV output, which is too large for the analyses generally conducted by the nominal climate model. Thus, a full sensitivity analysis including global and distributional sensitivity analysis was conducted. The results of these analyses are presented in the following subsection.

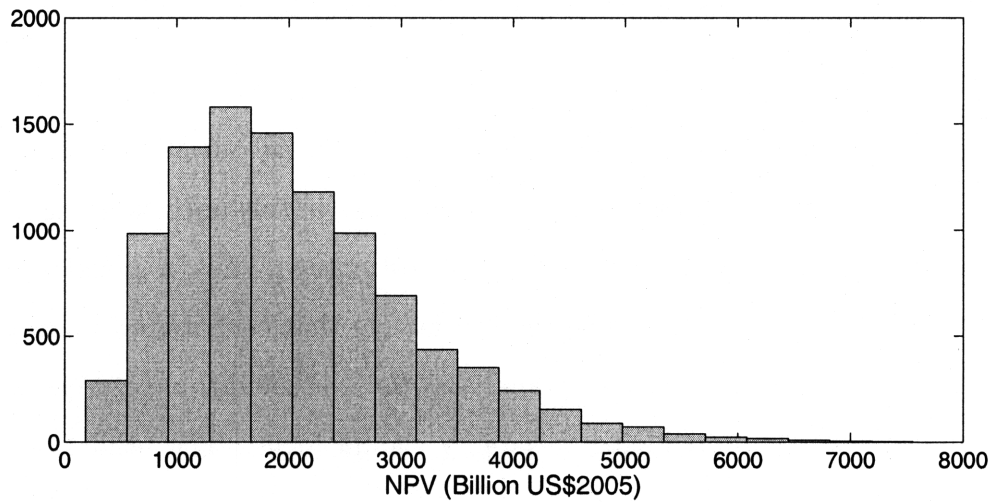


Figure 5-3: Histogram of Monte Carlo simulation results for NPV in billions of 2005 US dollars.

Since the principle of maximum entropy was employed to map the uncertainty information for each factor to a probability distribution, the output histogram shown in Figure 5-3 is the most likely distribution to be observed [41]. It is also important to point out that many studies favor a non-probabilistic approach to representing uncertainties such as those associated with the nominal climate model factors. Non-probabilistic approaches, such as interval analysis, are appealing because they do not make any assumptions regarding the likelihood of any events occurrence. For the case of the nominal climate model, the results of the sampling-based uncertainty propagation could be viewed from an interval analysis perspective as a range of possible outputs from about 200 billion to about 7500 billion 2005 US dollars, as shown by Figure 5-3. However, given that the damage coefficient is unbounded and NPV is a monotonic function of the damage coefficient, an analytical procedure for conducting interval analysis would result in a range of possible NPV values from negative infinity to positive infinity, which would not be a useful result.

## 5.2.6 Step 6: Conduct sensitivity analysis

The second question for furthering model development given in Section 5.2.1, is, “is there a need to direct research efforts aimed at reducing output variability.” As mentioned in the previous subsection, this need exists in this uncertainty assessment given the large standard deviation of the NPV estimates. Thus, a global sensitivity analysis was conducted to determine the key factors that contribute to output variability, and a distributional sensitivity analysis was conducted to determine on which factors research aimed at reducing variability should focus. The results of each of these analyses are presented below.

### Global sensitivity analysis of the nominal climate model

The main effect and total effect sensitivity indices of the NPV output of the nominal climate model are presented in Table 5.5. The indices are also displayed in bar chart form in Figures 5-4 and 5-5.

Table 5.5: Estimates of total and main effect sensitivity indices for the NPV output of the nominal climate model.

Factor	Main Effect	Total Effect
AEDT Fuelburn & CO <sub>2</sub> Emissions	0.01	0.01
AEDT NOx Emissions	0.00	0.00
Climate Sensitivity	0.32	0.39
Ref. Temperature Change	0.00	0.00
RF doubling of CO <sub>2</sub>	0.01	0.02
RF Short-lived Effects	0.12	0.13
Efficacies	0.00	0.00
Damage Coefficient $a_2$	0.44	0.52

The global sensitivity analysis results show that nearly all of the variability of the NPV output is due to the damage coefficient  $a_2$ , and the climate sensitivity factor. These factors were anticipated to be significant contributors to output variability, which provides confidence in the verification and validation of the model.

### Distributional sensitivity analysis of the nominal climate model

Since the variability of the NPV estimate is considered too large and thus must be reduced for the estimates to be useful, the next step in the sensitivity analysis process

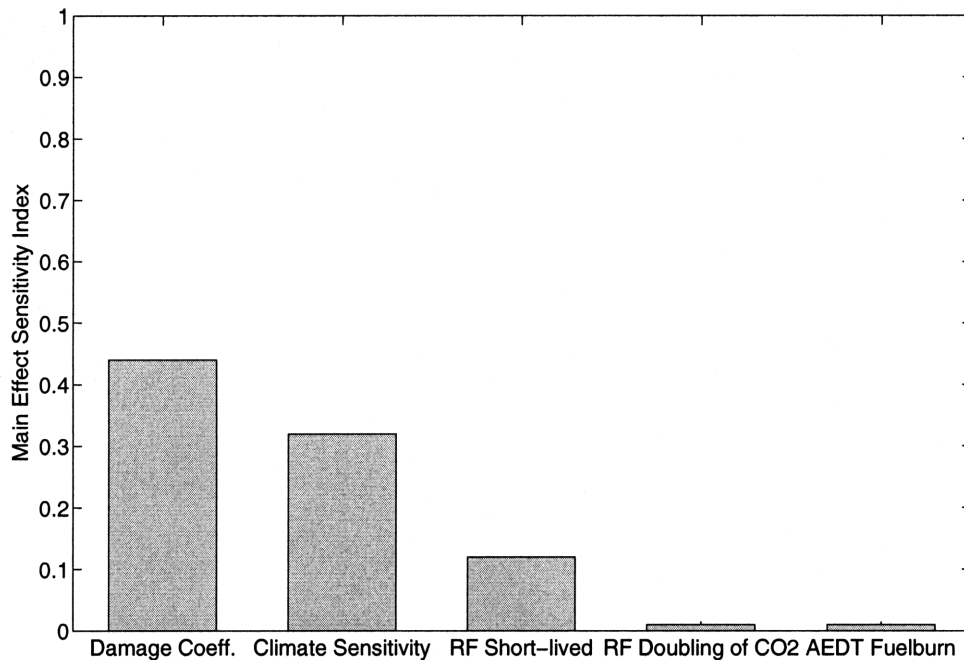


Figure 5-4: Main effect sensitivity indices of the NPV output of the nominal climate model.

is to determine on which factors research should focus to achieve a reduction in the variability of the NPV estimates through distributional sensitivity analysis. Given the main effect sensitivity indices from the global sensitivity analysis, the factors that should be considered are the damage coefficient, the climate sensitivity, and the radiative forcings of the short-lived effects, which combined, account for 88% of the variance of the NPV estimates. However, the nominal climate model has been developed in a manner that does not permit sensitivity analysis on the individual short-lived effects (the RF short-lived bar in Figure 5-4 represents the sum of the main effects of all of the short-lived effects), and thus, distributional sensitivity analysis cannot be performed on them individually. The short-lived main effects in total accounted for only 12% of output variability, and thus an analysis of only the damage coefficient and climate sensitivity would likely suffice, but this grouping of the effects is noted as a gap in the functionality of the model.

The adjusted main effect sensitivity indices for the damage coefficient and the

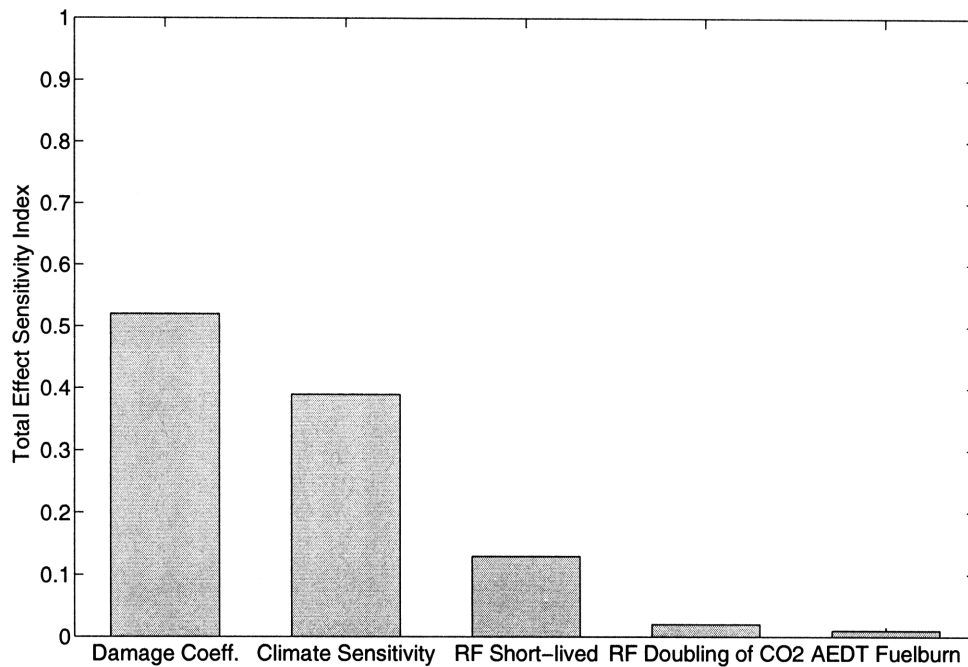


Figure 5-5: Total effect sensitivity indices of the NPV output of the nominal climate model.

climate sensitivity are presented in Figure 5-6. The average adjusted main effect sensitivity indices are given in Table 5.6.

As can be seen from the figure, the adjusted sensitivity indices for both the damage coefficient and climate sensitivity are close to linear in the amount of variance of each factor that can be reduced. This results in average adjusted main effect sensitivity indices that provide the same ranking as was provided by the main effect indices of global sensitivity analysis, however, the values of the average adjusted indices are lower by about a factor of 2, reflecting the fact that the reducible variance is considered a random variable in distributional sensitivity analysis, rather than assuming all variance is reducible as in global sensitivity analysis.



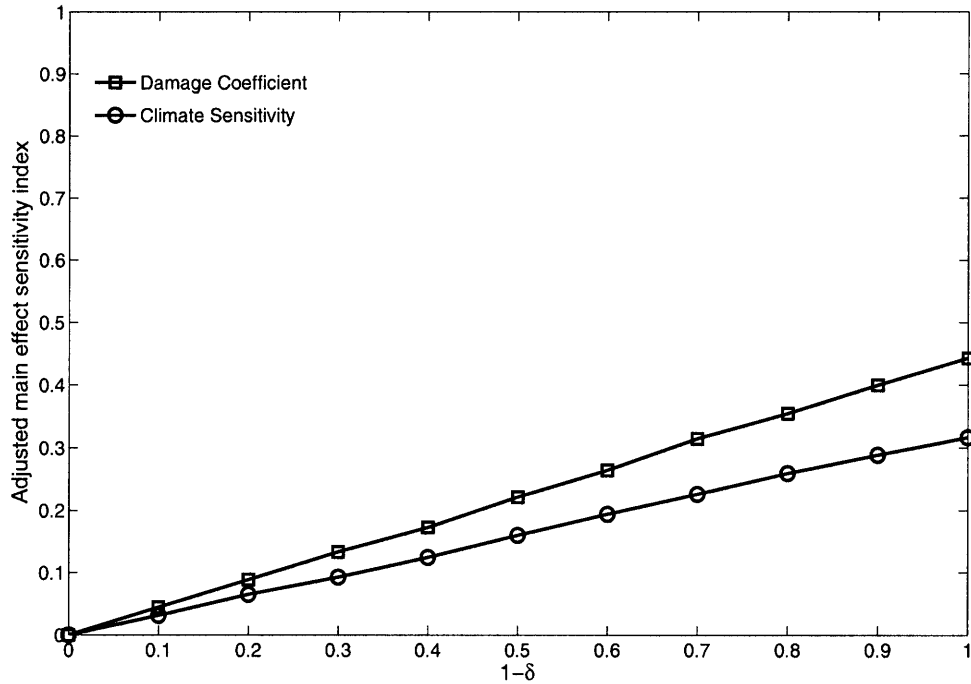


Figure 5-6: Adjusted main effect sensitivity indices of the NPV output of the nominal climate model.

Table 5.6: Estimates of averaged adjusted main effect sensitivity indices for the NPV output of the nominal climate model.

Factor	Average adjusted main effect sensitivity index
Damage Coefficient	0.22
Climate Sensitivity	0.16

### 5.3 APMT-Impacts Nominal Climate Model Uncertainty Assessment Conclusions

The uncertainty assessment of the nominal climate model presented here provides the information required to meet the goals set forth in Section 5.2.1. For decision-making, statistics such as the output mean and standard deviation were provided, as well as a histogram of output estimates, which can all be used in a quantitative evaluation of the performance of the model relative to fidelity requirements. Here, the estimated output variance is too large to support confidence in model results, which implied

the need for sensitivity analysis aimed at reducing output variability. The decision-making information gained in the uncertainty assessment could have also been used in a quantitative comparison of different policy scenarios, though comparisons of that nature were not made here.

For development, the results of the sensitivity analyses revealed that future development efforts aimed at trimming output variability should focus on the damage coefficient and climate sensitivity factors. Further, several gaps were identified in the process of performing the uncertainty assessment of the model. Among these were the inappropriate grouping of short-lived effects, as well as several instances of hard-coded values, such as the discount rate, that led to results that were incompatible with the model assumptions. These gaps raise the important point that for models that intend to include uncertainty assessment as part of model application and development, the model should be developed with the needs of the uncertainty assessment in mind. In this particular case, the discount rate parameter can be overwritten in several places in the climate module for analysis purposes unrelated to assessment, which can lead to critical mistakes in the uncertainty assessment of the models.

## **5.4 Further Uncertainty Assessment of the Climate Module**

The uncertainty assessment of the nominal climate model presented above, focused on the NPV output after 800 years. However, the climate module of the FAA environmental tools-suite is capable of looking at a variety of different climate related outputs after a variety of different time frames. One such output is the temperature change due to CO<sub>2</sub> emissions from aviation over a period of 5 years. This output is not necessarily a typical result of interest, however, sensitivity analyses of this particular output produce interesting results. Thus, in an effort to demonstrate the differences between global and distributional sensitivity analyses on a real-world model, these analyses are the topic of this section.

The factors that affect the temperature change due to CO<sub>2</sub> emissions from aviation are the climate sensitivity, the CO<sub>2</sub> emissions from aviation, and the RF doubling of CO<sub>2</sub>. Though for the nominal climate model climate sensitivity was assumed to be triangularly distributed with a minimum of 2, a maximum of 4.5, and a mode of 3, the source of the uncertainty information for that factor states that “large values of climate sensitivity cannot be excluded” [80] (though they generally are in analyses [3, 66, 67, 80]). Thus, a new distribution for the climate sensitivity is assumed, which is  $T(2, 10, 3)$ . The distributions for the CO<sub>2</sub> emissions and the RF doubling of CO<sub>2</sub> are as they were for the nominal climate model assessment. Using these distributions, both global and distributional sensitivity analyses were conducted for the temperature change due to CO<sub>2</sub> emissions from 5 years of aviation. The main effect sensitivity indices computed from global sensitivity analysis are shown in Figure 5-7. The figure shows that each factor has a substantial contribution to

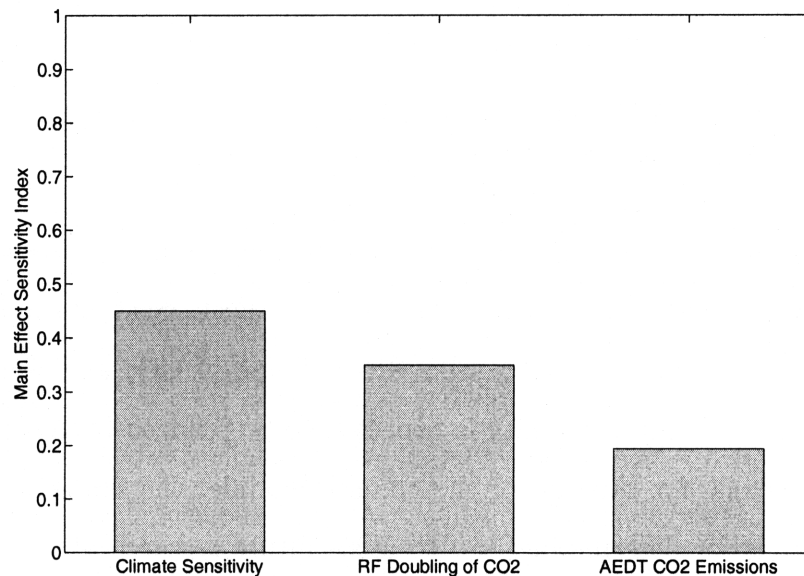


Figure 5-7: Main effect sensitivity indices for the temperature change due to CO<sub>2</sub> emissions from aviation after 5 years.

output variability, and that future research aimed at trimming output variability should focus on climate sensitivity first. However, as can be seen from the adjusted

main effect sensitivity indices computed using distributional sensitivity analysis in Figure 5-8, the factor on which future research should focus on may not be so obvious. Assuming the amount of variance that cannot be reduced through further research,

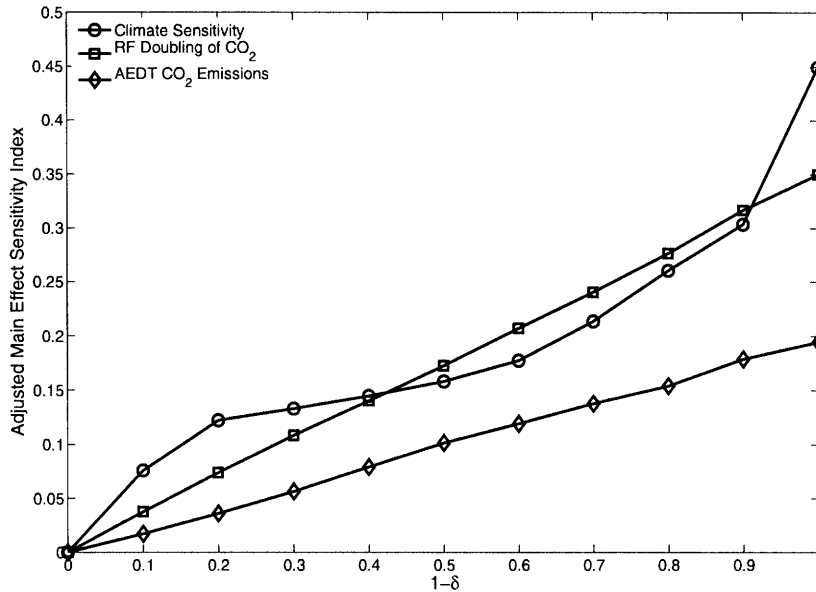


Figure 5-8: Adjusted main effect sensitivity indices for the temperature change due to CO<sub>2</sub> emissions from aviation after 5 years.

$\delta$ , is a uniform random variable on the interval  $[0,1]$ , the average adjusted main effect sensitivity indices, shown in Figure 5-9 with the main effect sensitivity indices of each factor, reveal that climate sensitivity and RF doubling of CO<sub>2</sub> are nearly equal in terms of how much output variability is expected to be reduced by researching each factor further. Thus, for this use of the climate module, the limitations of global sensitivity analysis for determining how to focus future research aimed at reducing output variability can lead to misleading conclusions, necessitating the application of distributional sensitivity analysis.

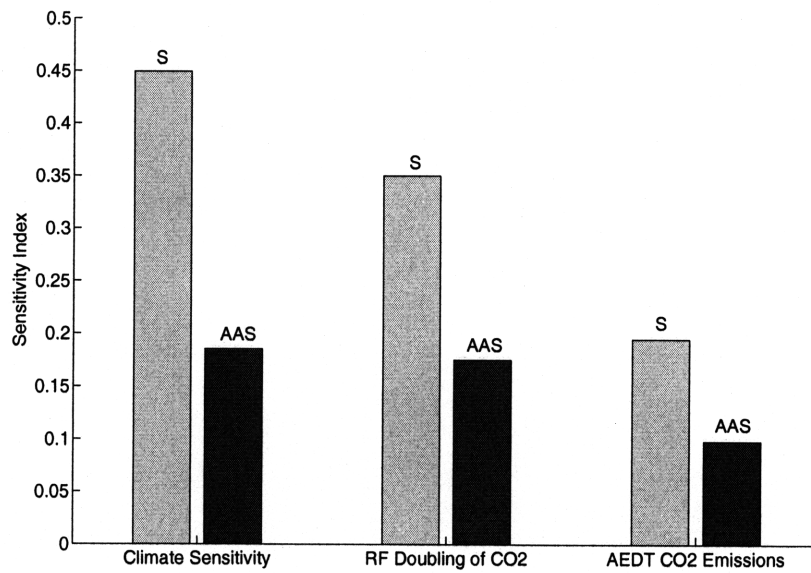


Figure 5-9: Averaged adjusted main effect and main effect sensitivity indices for the temperature change due to CO<sub>2</sub> emissions from aviation after 5 years.



# Chapter 6

## Conclusions and Future Work

The objectives of this research were threefold. They were to establish a probabilistic approach to uncertainty assessment for complex models intended to support decision-making and policy-making processes, to systematically develop surrogate models for situations where proper assessment of uncertainty is computationally prohibitive, and to demonstrate the application of the uncertainty assessment approach and surrogate modeling methodologies on real-world models. A summary of the work done to meet each objective is given in the following section, which is followed by general conclusions that can be drawn from this research, as well as a discussion of future research that should be considered.

### 6.1 Summary

A probabilistic approach to uncertainty assessment for complex models intended to support decision- and policy-making processes was established. The approach consists of seven steps: establishing goals, documenting assumptions and limitations, documenting factors and outputs, classifying and characterizing uncertainty, conducting uncertainty analysis, conducting sensitivity analysis, and presenting results. This approach meets the need for guidance in the application of uncertainty and sensitivity analysis on complex models.

A novel surrogate modeling methodology designed specifically for the uncertainty

assessment of an aviation environmental policy-making model was developed. The surrogate modeling methods developed allow for construction of rigorous confidence intervals for metrics that are useful for supporting decision-making (e.g. output means and variances), and for global sensitivity indices, which are useful for informing future research efforts aimed at furthering the development of a model, for a situation where running the analyses on a full model was infeasible. Furthermore, the methodology provides predictable convergence behavior of confidence interval widths from the surrogate model estimates, which allows for informed tradeoffs between computation time and uncertainty in the estimation of the various metrics.

The general approach to uncertainty assessment was demonstrated in detail on the APMT-Impacts nominal climate model. The demonstration showed that the probabilistic approach to uncertainty assessment is applicable to real-world models aimed at supporting decision- and policy-making processes. Further, the uncertainty assessment approach has been applied, or is currently being applied, to several models of the FAA environmental tools-suite, which gives confidence in the generality of the approach, and reveals the confidence real-world model developers have in the process.

## 6.2 Conclusions

As mentioned at the outset, numerical simulation is becoming increasingly widespread as a means to support decision-making and policy-making processes. The use of numerical simulation leads to such questions as: What confidence do we have in model results? What can be done to improve our confidence in model results? What are the limits in terms of applicability of our models to certain classes of problems? These important questions can all be rigorously answered by uncertainty assessment. Thus, uncertainty assessment is becoming an essential component of model development and application processes, and models should be developed with the needs of the uncertainty assessment in mind. As demonstrated in the assessment of the nominal



climate model, certain model functionalities such as hard-coding parameter values, which were implemented without consideration of uncertainty assessment needs, can potentially lead to incorrect conclusions from the uncertainty assessment and the wrong answers to our questions regarding confidence in model results.

An approach to uncertainty assessment using sampling-based probabilistic methods may lead to situations where analyses required for an uncertainty assessment are computationally prohibitive. For this reason, surrogate models can be essential components of an uncertainty assessment. In these situations, it is critical that uncertainty associated with the use of a surrogate model in the place of the full model be quantified. This process was demonstrated in this work for a particular class of aviation system models.

For complex models intended to support decision-making and policy-making processes, situations may also be encountered where our confidence in model results is not sufficient for using them in practice. In these situations, sensitivity analysis can be used to direct future development efforts aimed at improving our confidence in model results. For this application of sensitivity analysis, the newly-developed distributional sensitivity analysis can lead to different rankings than the state-of-the-art technique of global sensitivity analysis, which can lead to more informed decisions regarding how to focus future research.

## **6.3 Future Work**

Though the objectives of this research were met, the work has led to a number of new questions that should be considered in future work in this area. Regarding the general approach to uncertainty assessment, the use of other uncertainty models, such as possibility theory and Dempster-Shafer evidence theory should be considered given their ability to represent epistemic uncertainty. Within the approach presented here, different sampling strategies, such as latin hypercube and quasi-Monte Carlo should be explored, as well as techniques such as adaptive sampling and importance sampling, particularly in the area of acceptance/rejection sampling, which limited

this uncertainty assessment approach to brute force pseudorandom sampling.

Regarding the surrogate modeling strategy developed, a large number of model evaluations were used to compute operation-level  $\text{NO}_x$  emissions outputs in an effort to obtain accurate estimates of operation-level parameters so that uncertainty associated with those estimates could be neglected. However, it is the subject of future work to include the uncertainty in the estimation of operation-level parameters, which could lead to a smaller number of necessary model evaluations, and thus, even lower computational cost for performing the uncertainty and sensitivity analyses presented here.

Finally, application of the uncertainty assessment approach should be pursued on the models of the FAA environmental tools-suite to analyses where the outputs of interest are the differences between impacts on the environment under various scenarios. Further, the approach developed here should also be extended to guide the assessment of the full FAA tools-suite system, where such challenges as common factors across different models and correlation of model outputs that are used in downstream models will need to be addressed.

# Bibliography

- [1] U.S. FAA. U.S. FAA environmental tools suite assessment/evaluation effort, October 2007. ICAO Committee on Aviation Environmental Protection (CAEP) Modeling and Database Task Force (MODTF), CAEP/8-MODTF/2-IP/03, Bern, Switzerland.
- [2] B. Kim, G. Fleming, J. Lee, I. Waitz, J.P. Clarke, S. Balasubramanian, A. Malwitz, K. Klima, M. Locke, C. Holsclaw, L. Maurice, and M. Gupta. System for assessing aviation's global emissions (SAGE), Part 1: Model description and inventory results. *Transportation Research Part D: Transport and Environment*, 12(5):325–346, 2007.
- [3] K. Marais, S. Lukachko, M. Jun, A. Mahashabde, and I. Waitz. Assessing the impact of aviation on the climate. *Meteorologische Zeitschrift*, 17(2):157–172, 2008.
- [4] D. Cacuci. *Sensitivity and Uncertainty Analysis: Volume I, Theory*. Chapman & Hall/CRC, Boca Raton, Florida, 2003.
- [5] 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, Ltd., West Sussex, England, 2008.
- [6] J.C. Helton, J.D. Johnson, C.J. Sallaberry, and C.B. Storlie. Survey of sampling-based methods for uncertainty and sensitivity analysis. Sandia Report SAND2006-2901, June 2006.
- [7] J.P. Gwo, L.E. Toran, M.D. Morris, and G.V. Wilson. Subsurface stormflow modeling with sensitivity analysis using a Latin-hypercube sampling technique. *Ground Water*, 34:811–818, 1996.
- [8] H. Caswell, S. Brault, A. Read, and T. Smith. Harbor porpoise and fisheries: An uncertainty analysis of incidental mortality. *Ecological Applications*, 8:1226–1238, 1998.
- [9] M. Sanchez and S. Blower. Uncertainty and sensitivity analysis of the basic reproductive rate. *American Journal of Epidemiology*, 145:1127–1137, 1997.
- [10] *Merriam-Webster's Collegiate Dictionary*. Merriam-Webster, Incorporated, tenth edition, 1993. Definitions of *factor* and *parameter*, pages 416, 843.

- [11] A. Law and W. Kelton. *Simulation Modeling and Analysis*. McGraw-Hill, Inc., Boston, Massachusetts, 1982.
- [12] J. Sprient and G. Vansteenkiste. *Computer-Aided Modelling and Simulation*. Academic Press Inc. (London) Ltd., London; New York, 1982.
- [13] F. Neelamkavil. *Computer Simulation and Modelling*. John Wiley & Sons Ltd., Sussex, England, 1987.
- [14] J.C. Helton and F.J. Davis. Illustration of sampling-based methods for uncertainty and sensitivity analysis. *Risk Analysis*, 22:591–622, 2002.
- [15] A. Saltelli, T.H. Andres, and T. Homma. Sensitivity analysis of model output. *Computational Statistics and Data Analysis*, 15:445–460, 1993.
- [16] D.G. Cacuci and M.E. Schlesinger. On the application of the adjoint method of sensitivity analysis to problems in the atmospheric sciences. *Atmosfera*, 7:47–59, 1994.
- [17] K. Chan, A. Saltelli, and S. Tarantola. Sensitivity analysis of model output: Variance-based methods make the difference. Proceedings of the 1997 Winter Simulation Conference, 1997.
- [18] T. Homma and A. Saltelli. Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering and System Safety*, 52:1–17, 1996.
- [19] A. Saltelli and R. Bolado. An alternative way to compute Fourier amplitude sensitivity test (FAST). *Computational Statistics and Data Analysis*, 26:445–460, 1998.
- [20] B.M. Ayyub and G.J. Klir. *Uncertainty Modeling and Analysis in Engineering and the Sciences*. Taylor & Francis Group, Boca Raton, Florida, 2006.
- [21] I. Manno. *Introduction to the Monte-Carlo Method*. Akademiai Kiado, Budapest, 1999.
- [22] M. Kalos and P. Whitlock. *Monte Carlo Methods Volume 1: Basics*. John Wiley & Sons, Inc., New York, 1986.
- [23] L. Zadeh. Fuzzy sets. *Information and Control*, 8(3):338–353, 1965.
- [24] D. Dubois. Possibility theory and statistical reasoning. *Computational Statistics and Data Analysis*, 51:47–69, 2006.
- [25] S. Ciavatta, T. Lovato, M. Ratto, and R. Pastres. Global uncertainty and sensitivity analysis of a food web bioaccumulation model. *Environmental Toxicology and Chemistry: In Press*, 2008.

- [26] M.G. Morgan and M. Henrion. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge University Press, Cambridge; New York, 1990.
- [27] U.S. EPA. Guidelines for preparing economic analyses. EPA 240-R-00-003, September, 2000.
- [28] E. Hofer, M. Kloos, B. Kryzkacz-Hausmann, J. Peschke, and M. Woltereck. An approximate epistemic uncertainty analysis approach in the presence of epistemic and aleatory uncertainties. *Reliability Engineering and System Safety*, 77:229–238, 2002.
- [29] A. Ang and W. Tang. *Probability Concepts in Engineering: Emphasis on Applications to Civil and Environmental Engineering*. John Wiley & Sons, Inc., Hoboken, New Jersey, second edition, 2007.
- [30] Personal Communication with Dr. Andrea Saltelli, Venice, Italy, September 2008.
- [31] E.R. Tufte. *The Visual Display of Quantitative Information*. Graphics Press, Cheshire, Connecticut, 1983.
- [32] H. Ibrenk and M. Morgan. Graphical communication of uncertain quantities to nontechnical people. *Risk Analysis*, 7:519–529, 1987.
- [33] J.C. Helton. Conceptual and computational basis for the quantification of margins and uncertainty. Sandia National Laboratories Technical Report, in draft, SAND2009-XXXX, 2009.
- [34] R.L. Keeney and H. Raiffa. *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. John Wiley & Sons, New York, 1976.
- [35] R.L. Keeney. *Value-Focused Thinking*. Harvard University Press, Cambridge, Massachusetts, 1992.
- [36] R. de Neufville. *Applied Systems Analysis: Engineering Planning and Technology Management*. McGraw-Hill, New York, 1990.
- [37] T.M. Cover and J.A. Thomas. *Elements of Information Theory*. John Wiley & Sons, Inc., Hoboken, New Jersey, 1991.
- [38] T.M. Williams. Practical use of distributions in network analysis. *J. Opl Res. Soc.*, 43(3):265–270, 1992.
- [39] D. Johnson. The triangular distribution as a proxy for the beta distribution in risk analysis. *The Statistician*, 46(3):387–398, 1997.
- [40] 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 and System Safety*, 85:39–71, 2004.

- [41] E.T. Jaynes. *Probability Theory: The Logic of Science*. Cambridge University Press, Cambridge, United Kingdom; New York, 2003.
- [42] A. Mailhot and J.P. Villeneuve. Mean-value second-order uncertainty analysis method: Application to water quality modeling. *Advances in Water Resources*, 26:491–499, 2003.
- [43] Y. Jang, N. Sitar, and A. Der Kiureghian. Reliability analysis of contaminant transport in saturated porous media. *Water Resources Research*, 30(8):2435–2448, 1994.
- [44] S.K. Choi, R.V. Grandhi, R.A. Canfield, and C.L. Pettit. Polynomial chaos expansion with Latin hypercube sampling for estimating response variability. *AIAA Journal*, 42(6):1191–1198, 2004.
- [45] S. Isukapalli. *Uncertainty Analysis of Transport-Transformation Models*. PhD thesis, Rutgers, The State University of New Jersey, 1999.
- [46] G. Grimmet and D. Stirzaker. *Probability and Random Processes, 3rd Ed.* Oxford University Press, Oxford; New York, 2006.
- [47] H. Niederreiter. *Random Number Generation and Quasi-Monte Carlo Methods*. Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, 1992.
- [48] M.D. McKay, R.J. Beckman, and W.J. Conover. A comparison of three methods for selecting values of input variables in the analysis of outputs from a computer code. *Technometrics*, 21(2):239–245, 1979.
- [49] G. Marsaglia. The marsaglia random number cdrom with the diehard battery of tests of randomness. CDROM, 1995.
- [50] A. Saltelli, T.H. Andres, and T. Homma. Sensitivity analysis of model output: Performance of the iterated fractional factorial design method. *Computational Statistics & Data Analysis*, 20:387–407, 1995.
- [51] R.L. Iman, M.J. Shortencarier, and J.D. Johnson. A fortran 77 program and user's guide for the calculation of partial correlation and standardized regression coefficient. SANDIA National Laboratory Report SAND79-1472, 1980.
- [52] M.H. Kutner, C.J. Nachtsheim, J. Neter, and W. Li. *Applied Linear Statistical Models, Fifth Edition*. McGraw-Hill, Boston, Massachusetts, 2005.
- [53] I.M. Sobol'. Theorems and examples on high dimensional model representation. *Reliability Engineering and System Safety*, 79:187–193, 2003.
- [54] X. Wang. Improving the rejection sampling method in quasi-monte carlo methods. *Journal of Computational and Applied Mathematics*, 114:231–246, 2000.

- [55] M.H. DeGroot and M.J. Schervish. *Probability and Statistics, 3rd Ed.* Addison-Wesley, Boston, Massachusetts, 2002.
- [56] B. Efron. Better bootstrap confidence intervals. *Journal of the American Statistical Association*, 82(397):171–185, 1987.
- [57] R.P. Liem. Surrogate modeling for large-scale black-box systems. Master of Science thesis, Massachusetts Institute of Technology, 2007.
- [58] G. Venter, R. Haftka, and J.H. Starnes. Construction of response surface approximations for design optimization. *AIAA Journal*, 36(12):2242–2249, 1998.
- [59] A. C. Antoulas and D.C. Sorensen. *Approximation of Large-Scale Dynamical Systems*. SIAM, Philadelphia, Pennsylvania, 2005.
- [60] N. Alexandrov, R. Lewis, C. Gumbert, L. Green, and P. Newman. Approximation and model management in aerodynamic optimization with variable-fidelity models. *Journal of Aircraft*, 38(6):1093–1101, 2001.
- [61] T. Robinson, M. Eldred, K. Willcox, and R. Haimes. Strategies for multi-fidelity optimization with variable dimensional hierarchical models. AIAA paper 2006-1819, presented at the 2nd MDO Specialists Conference, Newport, RI, 2006.
- [62] S. Baughcomb, S. Henderson, and T. Tritz. Scheduled civil aircraft emissions inventories for 1976 and 1984: Database development and analysis. NASA CR-4722, 1996.
- [63] A. Karr. *Probability*. Springer, New York, 1993.
- [64] S. Lipschutz, J. Schiller, and R.A. Srinivasan. *Schaum's Outline of Theory and Problems of Beginning Finite Mathematics*. McGraw-Hill Professional, New York; London, 2004.
- [65] D. Horvitz and D. Thompson. A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, 47(260):663–685, 1952.
- [66] M. Jun. Uncertainty analysis of an aviation climate model and an aircraft price model for assessment of environmental effects. Master's thesis, Massachusetts Institute of Technology, 2007.
- [67] A. Mahashabde. *Comparative Analysis of Selected Policy Measures for Mitigating Aviation Environmental Impacts*. PhD thesis, Massachusetts Institute of Technology, 2009.
- [68] H.D. Jacoby. Informing climate policy given incommensurable benefits estimates. *Global Environmental Change Part A*, 14(3):287–297, 2004.

- [69] K. Hasselmann, S. Hasselmann, R. Giering, V. Ocana, and H.V. Storch. Sensitivity study of optimal CO<sub>2</sub> emission paths using a simplified structural integrated assessment model (SIAM). *Climatic Change*, 37(2):345–386, 1997.
- [70] R. Sausen and U. Schumann. Estimates of the climate response to aircraft CO<sub>2</sub> and NO<sub>x</sub> emissions scenarios. *Climatic Change*, 44(1-2):27–58, 2000.
- [71] J.S. Fuglestedt, T.K. Bernsten, O. Godal, R. Sausen, K.P. Shine, and T. Skodvin. Metrics of climate change: Assessing radiative forcing and emission indices. *Climatic Change*, 58(3):267–331, 2003.
- [72] K.P. Shine, J.S. Fuglestedt, K. Hailemariam, and N. Stuber. Alternatives to the global warming potential for comparing climate impacts of emissions of greenhouse gases. *Climatic Change*, 68(3):281–302, 2005.
- [73] W.D. Nordhaus and J. Boyer. *Warming the World: Economic Models of Global Warming*. MIT Press, Cambridge, Massachusetts, 2000.
- [74] G. Hooss, R. Voss, K. Hasselmann, E. Maier-Reimer, and F. Joos. A nonlinear impulse response model of the coupled carbon cycle climate system (niccs). *Climate Dynamics*, 18(3-4):189–202, 2001.
- [75] G. Plattner, F. Joos, T.F. Stocker, and O. Marchal. Feedback mechanisms and sensitivities of ocean carbon uptake under global warming. *Tellus*, 53B:564–592, 2001.
- [76] F. Joos, I.C. Prentice, S. Sitch, R. Meyer, G. Hooss, G. Plattner, S. Gerber, and K. Hasselmann. Global warming feedbacks on terrestrial carbon uptake under ipcc emission scenarios. *Global Biogeochemical Cycles*, 15:891–907, 2001.
- [77] INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE (IPCC). Special report on emissions scenarios. 2001.
- [78] INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE (IPCC). Climate change 2001: The scientific basis. In: IPCC Third Assessment Report.
- [79] D.S. Stevenson, R.M. Doherty, M.G. Sanderson, W.J. Collins, C.E. Johnson, and R.G. Derwent. Radiative forcing from aircraft NO<sub>x</sub> emissions: Mechanisms and seasonal dependence. *J. Geophys. Res.*, 109, 2004.
- [80] INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE (IPCC). Climate change 2007: Synthesis report. In: IPCC Fourth Assessment Report.
- [81] G. Yohe. Warming the world: Economic models of global warming. *J. Econ. Literature*, 40:573–574, 2002.
- [82] J. Hansen, M. Sato, and R. Ruedy. Efficacy of climate forcings. *J. Geophys. Res.-Atmos*, 110(D18), 2005.