Uncertainty Quantification and Confidence Assessment in Time-Dependent, Multidisciplinary Simulations

By

Erin Camille DeCarlo

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Civil Engineering

August 11, 2017

Nashville, Tennessee

Approved:

Sankaran Mahadevan, Ph.D.

Hiba Baroud, Ph.D.

Prodyot K. Basu, Ph.D.

Haoxiang Luo, Ph.D.

Benjamin P. Smarslok, Ph.D.

To Mom and Dad,

my sisters Ellen and Kelly

and

Grandaddy

ACKNOWLEDGEMENTS

Reflecting on the past six years I have spent at Vanderbilt, I am overwhelmed with gratitude and thankfulness for having had the opportunity to work with and learn from so many people. First, I want to thank my advisor, Dr. Maha for seeing potential in me and supporting me throughout the process of earning my Ph.D. I am in awe of his dedication and sacrifice to his students, and consider myself lucky to be counted among them. His example as a researcher and leader is one I can only strive for and he is truly a father figure to so many of us. I can only hope the summation of all our dissertation acknowledgements do him justice.

Next, I want to thank my AFRL mentor, Ben Smarslok. I value the time I have spent working with him and the folks in the Structural Sciences Center, and I am thankful to have been welcomed into their research family. Ben both funded my research directly and helped me obtain a National Defense Science and Engineering Graduate (NDSEG) research fellowship, but more importantly, I am grateful to him for his mentorship and friendship. I can recall countless instances that he fought for me when he did not have to, mostly when I did not think I deserved it, and this document is a testament to all the work he has done. Truly, this dissertation would not have been remotely possible without him.

Finally, I want to thank my family – my parents, Chris and Angie, and my sisters, Ellen and Kelly – who have supported me at every stage of my education. I also want to thank all the friends I have made at Vanderbilt – especially T, Lyndsey, Chelsea, Nate, Josh, Yonathan, Bethany, Chen, Amy, Janelle, and Joe – for my sanity and all the best trivia, karaoke, and game nights I could ask for. Also, I am thankful for the friends I met through AFRL – Daniel, Diane, Carolina, and Rob – for so many good memories in Dayton as well as their unique insights and contributions into my research. These are all relationships and friendships I will hold dear for the rest of my life.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

CDF	Cumulative density function
DBN	Dynamic Bayesian network
DL	Double loop
FTSI	Fluid-thermal-structural interactions
GP	Gaussian process
GSA	Global Sensitivity Analysis
HTT	High-Temperature Tunnel
ISK	Importance sampling-based kernel regression
KOH	Kennedy and O'Hagan
LHS	Latin-hypercube sampling
MCE	Matrix-column exchange
MCS	Monte Carlo simulation
MDA	Multidisciplinary analysis
MDO	Multidisciplinary optimization
NASA	National Aeronautics and Space Administration
PDF	Probability density function
QoI	Quantity of interest
TPS	Thermal protection system
USAF	United States Air Force
UQ	Uncertainty quantification

CHAPTER 1

INTRODUCTION

1.1 Motivation: Reusable Hypersonic Vehicles and Digital Aircraft Twin

The U.S. Air Force (USAF) has identified the need for reusable high-speed aircraft that can endure the combined aerodynamic, thermal, and acoustic loads observed in hypersonic flight environments [1]. Due to limited operational data and the inability to fully reproduce hypersonic loads in ground facilities, there is significant uncertainty surrounding the structural response and material degradation of an aircraft structure that is repeatedly subjected to fluctuating aero-pressures and high thermal-gradients from high-Mach flights [2]. As a result, the highly-coupled aerothermoelastic response of a structure characterized by fluid-thermal-structural interactions (FTSI) is studied primarily via single- and partial-physics constituent models (e.g., aero-elastic, fluid-thermal, heat transfer, aero-acoustic, etc.) [3,4].

Concurrently with the increased emphasis on computer simulations to study the aerothermoelastic response, the USAF has identified the need to create a multi-physics simulation model to predict the damage, life, and reliability of a vehicle prior to and throughout its assigned mission [2,5]. This model is referred to as an "Aircraft Digital Twin," which is a continuously updating computer simulation that makes mission decisions and adjustments based on on-board in-flight measurements, manufacturing and machining variations, and the inspection, repair, and operation history of an individual aircraft [6]. However, full-fidelity simulations using high spatial, temporal, and coupling resolutions for each multidisciplinary model component and interaction are computationally expensive [7]. Thus, there is ongoing research focused on improving computational efficiency by using simplified physics, surrogates, and reduced-order models for FTSI simulations [8–11]. To that end, multi-level, multi-disciplinary, and multi-fidelity models

are to be integrated into an uncertainty-informed decision-making framework to streamline vehicle certification, optimize fleet management, and facilitate sustainment [12].

In the absence of full-scale physical tests under fully representative hypersonic loads, research is underway to fill-in knowledge gaps with small-scale component-level tests (e.g., panel) that capture subsets of the relevant physical interactions [13–17]. Historically, there has been very limited data from FTSI experiments in high-speed flows. A few of the validation-quality aerothermal loading experiments considered rigid flat plates, spherical domes, and quilted domes and were conducted at the NASA 8ft High-Temperature Tunnel (HTT) by Deveikis et al [18,19] and Glass and Hunt [20,21] during the 1970s and 1980s. In particular, the Glass and Hunt experiments in 1986 [20] were designed to study variable surface pressure and heating rates across rigid spherical domes in laminar and turbulent Mach 6.5 flow, and provide aerothermal data for the heating-induced deformations expected to occur during flight of metallic TPS panels. These legacy aerothermal HTT tests do not capture the structural dynamics that would constitute FTSI, however, they have supplied benchmark pressure and heating profiles for several static panel configurations to compare against the fluid-thermal components in FTSI solvers [22–24]. Thus, these limited, heterogeneous, and partial information sources can be used to quantify errors and uncertainty in their respective sub-domains of aerothermoelasticity, and help guide new experiments toward reducing significant knowledge gaps surrounding FTSI [25].

Due to limited wind tunnel data and prohibitively expensive flight tests, decision-making relies heavily on uncertainty-quantified predictive models. This dissertation investigates effective uncertainty quantification methods for multidisciplinary, transient models (i.e., model calibration, model confidence assessment, global sensitivity analysis, and model selection) with specific focus

toward reducing the uncertainty and improving the accuracy and efficiency of coupled, aerothermoelastic predictions.

1.2 Uncertainty in Aerothermoelastic Simulations

Uncertainty inherently exists due to imperfect data and knowledge which includes: 1) natural variability in the system and model inputs (e.g., materials, geometry, fluid properties, and loading); 2) uncertainty from measurement errors and limited data; and 3) and model uncertainty and errors from simplified or poorly-understood physics, their interactions, and numerical approximations [26]. At high Mach numbers, fluid-structural coupling between aerodynamic pressure (p) and a deforming structure (w) is accompanied by significant aerodynamic heat flux (Q_{aero}) from large temperature gradients within the fluid boundary layer [27]. These aerodynamic heat loads transfer through the structure, augment the temperature-dependent material properties, and lead to non-uniform thermal gradients in $T_{structure}$ which cause further structural deformation into the flow. A monolithic solution for the aerothermoelastic problem would consist of simultaneously solving the unsteady Navier-Stokes equations with real gas effects and structural dynamics equations. Due to the computational burden of a transient monolithic solution over a full aircraft, partitioned aero-thermal-structural coupling strategies like the one shown in Figure 1.1 have been pursued to simulate and study FTSI [28–31].

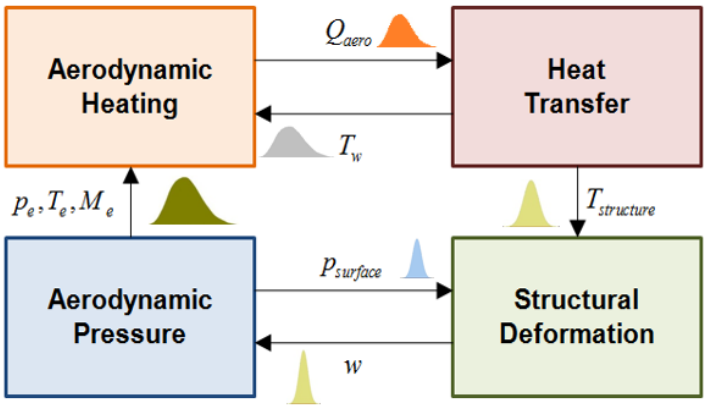

Figure 1.1. Coupling in aerothermoelasticity

Partitioned approaches lend themselves to modularization and simulation flexibility where under a constrained computational budget (e.g., combat mission planning with a Digital Twin) either a) simplified physics or reduced-order models can be substituted for an expensive component, or b) the coupling strength between different disciplinary models can increase or decrease based on the application domain. In aerothermoelastic modeling, two viable alternatives have been identified: 1) simplified representations of the unsteady aerodynamic pressure (i.e., 3rd-order piston theory [32]) and heating loads (i.e., Eckert's reference temperature method [33]); and 2) efficient incorporation of heat transfer between the fluid and structure into an aeroelastic solution process [34,35]. However, each model and coupling approximation contain errors and uncertainty, and also affect how the aforementioned sources of uncertainty propagate across the coupling interfaces.

Model uncertainty impedes the realization of reusable hypersonic vehicles and the Aircraft Digital Twin due to the limited ability to perform ground tests and confidently predict the structural reliability, risk, and performance of these aircraft [36]. To make progress toward these long-term goals, this dissertation investigates effective uncertainty quantification methods for multidisciplinary, transient simulation models (i.e., model calibration, model confidence

assessment, global sensitivity analysis, and model selection) with the goal of improving the accuracy and efficiency of coupled, aerothermoelastic simulation.

1.3 Outline of Dissertation

This dissertation focuses on quantifying the confidence in multidisciplinary simulations, which are often assembled based on limited data and inadequate individual and partial-physics model components. Methodologies are proposed to address the following three challenges in the context of both inverse and forward uncertainty quantification (UQ) problems: 1) computational expense of multidisciplinary simulations, 2) error accumulation across multiple models and over time, and 3) uncertainty due to the availability of limited data.

The current state of the art in UQ analysis mostly addresses single-physics problems. Very few studies consider multidisciplinary analysis where errors and uncertainty aggregate across disciplinary models [37] (only forward prediction, not inverse problems). Therefore, this dissertation aims to fill this gap with respect to multidisciplinary model calibration and prediction. Specifically, regarding inverse problems, the interest lies in isolating model error contributions and for prediction problems, the goal is to improve prediction confidence and minimize computational effort.

For inverse problems, the development of a segmented Bayesian model calibration strategy reduces the computational effort of calibration when multiple information sources are available. Further, prediction confidence is improved by reducing the uncertainty that aggregates between coupled analyses and through time using a partitioned approach to calibrate model errors. Methodology contributions for the forward problem include an efficient global sensitivity analysis method (to support dimension reduction) that incorporates existing model calibration results and an optimization framework that balances prediction confidence and computational effort to select

variable model fidelity in multidisciplinary simulations. These methods are illustrated with time-dependent, aerothermoelastic analyses of airfoils subjected to high-speed flow.

Objective 1: Bayesian calibration strategies for multidisciplinary models

Bayesian calibration methods are explored to improve confidence in coupled aeropressure and aeroheating predictions using data on multiple quantities of interest. First, models and data are systematically integrated using a Bayesian network approach, and the objective is to investigate that impact of an alternative segmented calibration strategy compared to traditional simultaneous calibration over the entire network. The calibrations are performed using historic aerothermal wind tunnel test data with turbulent Mach 6.6 flow [20] and evaluated on the basis of increased prediction confidence and computational effort. The segmented calibration strategy targets dominant connections between uncertainty sources in the presence of limited data for maximum uncertainty reduction.

Objective 2: Model discrepancy calibration in time-dependent, coupled analyses

This Bayesian network framework from Objective 1 is extended to include transient heat transfer and time-series temperature data for the calibration of model discrepancy in the coupled aerodynamic heating and heat transfer prediction. The objective is to adapt the dynamic Bayesian network for time-dependent aerothermal models using the appropriate model discrepancy resolution for isolation and propagation of uncertainty through the network. Global, step-wise, and partitioned approaches for inferring coupled model errors are compared, where the partitioned approach is shown to capture the nonlinear response in extrapolation with minimal effect on the calibration cost and no effect on the cost of forward prediction.

Objective 3: Efficient global sensitivity analysis for transient, multidisciplinary problems

The objective is to improve the efficiency of the global sensitivity analyses (GSA) that occur both before and after the model calibrations in Objectives 1 and 2. For pre-calibration GSA, convergence of an importance-sampling based global sensitivity analysis methodology (ISK-GSA) [38] is improved with low-discrepancy sampling sequences (i.e., Sobol' sequence), thus minimizing the number of pre-calibration simulations required. For post-calibration GSA, the ISK-GSA methodology is generalized to efficiently compute global sensitivities using a) existing input-output relationships from calibration; and b) correlated input variables.

Objective 4: Model selection and coupling in time-dependent, multidisciplinary simulations

The final objective is to develop a confidence-based model selection framework to determine optimal coupling and temporal fidelities while accounting for uncertainty and error propagation differences among the simulation alternatives. At each time-step, an optimization decides the next model coupling and time-step size needed to balance both prediction confidence and simulation cost constraints. The model-selection process is demonstrated on the calibrated aerodynamic heating and heat transfer models from Objective 2.

The above four dissertation objectives are pursued in Chapters 2 – 5 in the order above. Finally, Chapter 6 presents a summary of the research contributions and directions for future work.

CHAPTER 2

BAYESIAN CALIBRATION OF MULTIDISCIPLINARY MODELS

2.1 Introduction

Within the probabilistic risk assessment and structural reliability communities, sources of uncertainty are characterized as either *aleatory* (irreducible) or *epistemic* (reducible) [39]. Aleatory sources of uncertainty are treated as random variables in the system with a probability density function (PDF) of known form (e.g., physical variability and measurement errors). Reliability assessments are then performed by propagating these aleatory uncertainty sources through mathematical representations (i.e. models) of the physical system using techniques such as Monte Carlo simulation (MCS), the first-order reliability method (FORM), and the second-order reliability method (SORM) [40]. Epistemic sources of uncertainty, however, arise from lack of knowledge about the physical system (e.g., model form errors, limited data) and affect the downstream reliability prediction. In the Bayesian approach, these sources of uncertainty are also represented through PDFs but are assumed to have a true deterministic value that is unknown. The uncertainty about their true value can be reduced using Bayesian updating techniques upon acquiring new information.

The goal of model calibration, therefore, is to use the available data to reduce the uncertainty regarding the sources of epistemic uncertainty that affect the quantity of interest (QoI), which may be a model prediction, prediction confidence, or the reliability of the system predicted by one or more models [41]. Several model calibration and parameter estimation methodologies exist (e.g., least squares, maximum likelihood estimation (MLE), and Bayesian estimation). Within a Bayesian calibration framework, knowledge on the uncertain parameters both prior to and after calibration are represented as prior and posterior PDFs which can be used for post-calibration

uncertainty propagation, model confidence assessment, and reliability analyses [42,43]. Therefore, Bayesian approaches are pursued for the coupled aerothermoelastic simulation where the overall goal is to quantify and increase model confidence in each of the models corresponding the disciplines in Figure 1.1.

When applying Bayesian methods over large systems for overall uncertainty reduction, it is necessary to connect the available data and the sources of uncertainty in a systematic way. Bayesian networks can be used for this purpose where the model inputs, outputs, data, and uncertainty sources are represented as nodes in the Bayesian network. The value of Bayesian networks lies in their ability to reduce the uncertainty over the entire network even when limited data is observed on even a small number of nodes. Bayesian networks and have been used for multi-level and hierarchical systems [44–46], where data is available at multiple levels (e.g., material, component, assembly). In these types of systems, each level feeds into the next and higher-level data can reduce the uncertainty regarding lower-level parameters. Bayesian calibration studies for these types of systems have compared the effects of calibrating low-level parameters with both lower-level and higher-level data, and have developed a roll-up technique to combine the posteriors from these calibrations using weights derived from single-level model validations [26,47]. Mullins et al [48] and Li and Mahadevan [46] further demonstrated the role of calibration and validation data available on each level in the network and optimized data collection to maximize the reliability of the top-level prediction within the roll-up framework.

Similarly to multi-level and hierarchical Bayesian networks, multidisciplinary Bayesian networks have nodes representing the simulation components and uncertainty sources from each discipline. Each disciplinary model output is connected to one or more downstream models as inputs or the analysis in next time instant within the same discipline. (When connecting two time

instants, the Bayesian network is constructed as a dynamic Bayesian network (DBN), which is discussed in Chapter 2). The differences between multi-level and multidisciplinary Bayesian networks, however, lie in the heterogeneity of data available on a single model component. For example, single-physics experiments may be available as well as combined-physics experiments that capture cross-disciplinary interactions. Within the multidisciplinary Bayesian network, combined-physics experiments contain information on multiple nodes in the network, much like higher-level data in multi-level systems; however, updating simultaneously over the all the relevant nodes and data in the system may be impractical, due to a) the uncertainty being sufficiently reduced from previous calibration with a single-physics data set, or b) one or more of the models within relevant disciplines being computationally cumbersome. Therefore, calibration priority should be given to the uncertainty sources that will most benefit from the data.

The segmented Bayesian calibration strategy developed in this chapter exploits the dominant connections between data and key sources of uncertainty in the presence of both single- and combined-physics data. The segmented calibration process partitions the parameter space and leads to greater uncertainty reduction within the calibration segments. When new data becomes available, the segmented Bayesian calibration strategy makes efficient use of the available data and reduces the computational burden by sequentially updating only the relevant calibration parameters and models. This motivates the comparison of segmented and simultaneous calibration procedures. Two comparison metrics are used in this study: the Bayes factor metric [49] is used to assess prediction accuracy, and the Kullback-Leibler (K-L) divergence metric [50] is used to compare the convergence rates of the posterior distributions, which are indicative of the overall computational effort of implementing the chosen calibration strategy.

Bayesian model calibration is presented in Section 2.2 and the Bayes factor and Kullback-Leibler divergence metrics used to compare the calibration strategies are presented in Section 2.3 and Section 2.4, respectively. The segmented and simultaneous calibration strategies for Bayesian model calibration are presented in Section 2.5 and demonstrated on illustrative examples in Section 2.6 and the aeropressure and aeroheating models with HTT wind tunnel data in Section 2.7.

One additional outcome of this chapter is the identification of the key model dependence and data characteristics for which a segmented Bayesian calibration strategy offers the most computational benefit without compromising downstream prediction reliability. The characteristics considered are: (1) the degree of dependence between models; (2) the relative numbers of single and combined effect experiments; and (3) the presence of shared parameters. Appendix A contains the analytical derivations of the posterior distributions from both segmented and simultaneous calibrations of a linear two-parameter example, where each parameter represents the uncertainty contribution from each physics. These derivations offer first-order estimates of the prediction reliability after a) single and combined-effect experiments, and b) segmented or simultaneous calibration, and can be readily implemented in resource allocation frameworks for both design of calibration experiments and model updating.

2.2 Bayesian Model Calibration Methodology

Consider a quantity of interest y predicted by a single physics model $\mathcal{M}_y$ that maps inputs $\boldsymbol{x}$ and model parameters $\boldsymbol{\theta}$ to the model prediction $\hat{y}$, which is an inexact estimate of the true value of y. The difference between available observations y_D and the true value of y is attributed to measurement error ε_D, as shown in Eq. (2.1) and often treated as a Gaussian random variable with zero mean and unknown variance σ_D^2. To capture the disagreement between $\hat{y} = \mathcal{M}_y$ and y due to

missing physics or approximations in the model $\mathcal{M}_y$, a model discrepancy term $\delta(\boldsymbol{x})$ parameterized as a function of the inputs is introduced in Eq. (2.2).

$$y_D = y + \varepsilon_D \tag{2.1}$$

$$y = \mathcal{M}_y(\boldsymbol{x};\boldsymbol{\theta}) + \delta(\boldsymbol{x}) \tag{2.2}$$

Input variables $\boldsymbol{x}$ are measurable quantities in laboratory or field experiments. These could be considered deterministic or stochastic with known probability distributions due to measurement errors or natural variability. In contrast to $\boldsymbol{x}$, model parameters $\boldsymbol{\theta}$ are uncertain due to lack of knowledge. Furthermore, the precise relationship between $\delta(\boldsymbol{x})$ and $\hat{y}$ is unknown. Thus, the goal of Bayesian model calibration is to use data y_D to estimate the posterior distributions of the parameters $\boldsymbol{\theta}$, $\delta(x)$, and σ_D^2, given observations from physical experiments and assumed prior distributions of these parameters. In some problems, some of the unmeasured inputs may also be treated similar to calibration parameters. Epistemic uncertainty may be present regarding either deterministic inputs (i.e., unknown value), or stochastic inputs (i.e., unknown distribution type and/or parameters). In this dissertation objective, epistemic uncertainty in the input parameters is not explicitly considered since its treatment is similar to model parameters $\boldsymbol{\theta}$.

For a single computational model, Figure 2.2 shows the Bayesian network corresponding to Eqs. (2.1) and (2.2). Bayesian calibration is an inverse problem, which is achieved by passing the information upstream from the data nodes (solid squares) to the calibration quantities (dashed circles).

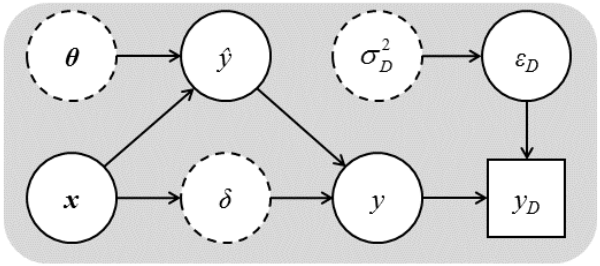

Figure 2.1. Bayesian network for a single physics model $\mathcal{M}_y$

Bayesian model calibration for continuous variables is facilitated using Bayes' formula shown in Eq. (2.3) where the posterior probability density of the calibration quantities $\mathbf{\Phi} = [\theta, \delta(x), \sigma_D^2]$ is proportional to the product of the joint likelihood $L(\mathbf{\Phi})$ and the prior probability density $\pi(\mathbf{\Phi})$. Here, $\pi(\bullet)$ denotes a probability density function (PDF) and $\pi(\mathbf{\Phi}|\, y_D)$ refers to the updated PDF (i.e., posterior) of parameters $\mathbf{\Phi}$ after observing y_D.

$$\pi(\mathbf{\Phi} \mid y_D) = \frac{L(\mathbf{\Phi})\pi(\mathbf{\Phi})}{\int L(\mathbf{\Phi})\pi(\mathbf{\Phi})d\mathbf{\Phi}} \propto L(\mathbf{\Phi})\pi(\mathbf{\Phi}) \tag{2.3}$$

Given N independent observations of y_D the joint likelihood function of parameters $\mathbf{\Phi}$ is shown in Eq. (2.4). The likelihood is proportional to the product of joint conditional probabilities of observing y_D given a deterministic value of $\mathbf{\Phi}$.

$$L(\mathbf{\Phi}) \propto \prod_{i=1}^{N} \pi(y_{D,i}, \mathbf{\Phi}) \propto \pi(y_D \mid \mathbf{\Phi}) \tag{2.4}$$

The integral in the denominator of Eq. (2.3) simply normalizes the posterior distribution to a valid PDF, however, numerical integration schemes quickly become intractable with increasing dimension of $\mathbf{\Phi}$ [51]. For this reason, Bayesian calibration is often performed using Markov chain Monte Carlo (MCMC) sampling algorithms – such as Metropolis-Hastings [52], Gibbs [53], or slice sampling [54] – which make direct use of the proportionality between the numerator $L(\mathbf{\Phi})\pi(\mathbf{\Phi})$ and the posterior distribution from which samples desired. Slice sampling is used in this

dissertation. Furthermore, the cost of computing the likelihood function is proportional to the cost of the model being calibrated, and can be demanding when costly models are used to calibrate numerous parameters. These challenges are discussed in more detail in Section 2.5 in the context of multidisciplinary Bayesian model calibration when one or more models may be computationally prohibitive.

2.3 Bayes Factor Metric

For models whose parameters are estimated through Bayesian methods and the posterior parameters $\mathbf{\Phi}$ are stochastic, the likelihood ratio of observing the data (known as Bayes factor) can be employed for selecting between two competing models or methods, $\mathcal{M}_1$ and $\mathcal{M}_2$. For example, the posterior of parameters $\mathbf{\Phi}_i$ given model $\mathcal{M}_i$ and calibration data D_c is shown in Eq. (2.5) as a model-dependent extension of Eq. (2.3)

$$\Pr(\mathbf{\Phi}_i \mid D_c, \mathcal{M}_i) = \frac{\Pr(D_c \mid \mathbf{\Phi}_i, \mathcal{M}_i) P(\mathbf{\Phi}_i)}{\int \Pr(D_c \mid \mathbf{\Phi}_i, \mathcal{M}_i) P(\mathbf{\Phi}_i) d\mathbf{\Phi}_i} \tag{2.5}$$

Often a validation set of data D_v is held separately from D_c to evaluate the predictive capabilities of the calibrated model $\mathcal{M}_i$ which is defined as $\Pr(D_v \mid \mathbf{\Phi}_i, \mathcal{M}_i)$, the likelihood of observing D_v given $\mathcal{M}_i$ and parameters $\mathbf{\Phi}_i$. The ratio of each $\Pr(D_v \mid \mathbf{\Phi}_i, \mathcal{M}_i)$ for each $i = 1$ to 2 defines the Bayes factor between two alternatives, shown in Eq (2.6). Since the parameters $\mathbf{\Phi}_1$ and $\mathbf{\Phi}_2$ are not deterministic but defined according to the posterior in Eq. (2.5), the integral over the posteriors of $\mathbf{\Phi}_1$ and $\mathbf{\Phi}_2$ defines the Bayes factor $B(\mathcal{M}_1, \mathcal{M}_2)$, where $B > 1$ indicates that the validation data D_v supports the posterior distribution derived from calibrating $\mathcal{M}_1$. Inherently, the Bayes factor accounts for limited data in the calibration of $\mathbf{\Phi}_i$ and when the maximum *a posteriori* estimate of $\mathbf{\Phi}_i$ is used, the Bayes factor is equivalent to a likelihood ratio test.

$$B(M_1, M_2) = \frac{\Pr(D_v \mid \mathcal{M}_1)}{\Pr(D_v \mid \mathcal{M}_2)} \equiv \frac{\int \Pr(D_v \mid \boldsymbol{\Phi}_1, \mathcal{M}_1) \Pr(\boldsymbol{\Phi}_1 \mid D_c, \mathcal{M}_1) d\boldsymbol{\Phi}_1}{\int \Pr(D_v \mid \boldsymbol{\Phi}_1, \mathcal{M}_2) \Pr(\boldsymbol{\Phi}_2 \mid D_c, \mathcal{M}_2) d\boldsymbol{\Phi}_2} \tag{2.6}$$

In this chapter, the Bayes factor in Eq. (2.6) is compares posteriors from two competing Bayesian model calibration strategies – segmented and simultaneous calibration with posteriors $\boldsymbol{\Phi}_{seg}$ and $\boldsymbol{\Phi}_{sim}$ respectively– that are presented in Section 2.5, tested on a mathematical example problem in Section 2.6, and on interacting aeropressure and aeroheating models in Section 2.7.

2.4 Kullback-Leibler Divergence

Computational effort is compared using the posterior convergence rates from both segmented and simultaneous calibration strategies, which can be computed using the Kullback-Leibler (K-L) divergence measure presented in Eq. (2.7) [50]. The K-L divergence is an indicator of similarity between PDFs, so a smaller value of K-L divergence indicates a smaller dissimilarity between them. The distributions of interest are the posterior distributions obtained from the calibration after the i^{th} MCMC sample $\pi_i(\boldsymbol{\Phi})$ and the posterior distributions after $i+1$ samples, $\pi_{i+1}(\boldsymbol{\Phi})$. The K-L divergence integral in Eq. (2.7) is computed using Monte Carlo integration with the samples generated from by the MCMC algorithm. Slice sampling is used here, but other MCMC methods could also be used.

$$D_{KL}(\pi_i \,||\, \pi_{i+1}) = \int \pi_i(\boldsymbol{\Phi}) \log \frac{\pi_i(\boldsymbol{\Phi})}{\pi_{i+1}(\boldsymbol{\Phi})} d\boldsymbol{\Phi} \geq 0 \tag{2.7}$$

Therefore, a simultaneous calibration strategy will track the K-L divergence of the posterior PDF of $\boldsymbol{\Phi}$ until convergence is reached. In contrast, a segmented calibration strategy monitors the K-L divergence of each calibration segment. For the illustrative examples in Section 2.6 and the hypersonic application problem in Section 2.7, two calibration segments are used.

2.5 Strategies for Multidisciplinary Bayesian Model Calibration

Consider two physical model predictions $\hat{y}_1(\boldsymbol{x}; \boldsymbol{\theta}_1, \boldsymbol{\theta}_{12})$ and $\hat{y}_2(\boldsymbol{x}, y_1; \boldsymbol{\theta}_2, \boldsymbol{\theta}_{12})$ which are estimates of y_1 and y_2, respectively. Each model contains inputs $\boldsymbol{x}$, uncertain model parameters $\boldsymbol{\theta}_i$, and input-dependent discrepancy functions $\delta_i(\boldsymbol{x}, y_{i-1})$. There are additional uncertain parameters common to both models represented by $\boldsymbol{\theta}_{12}$. Building off of the notation in Eq. (2.2), the mathematical relationships for this two-discipline system are defined in terms of their model outputs, uncertain parameters, and model discrepancy in Eqs. (2.8) and (2.9).

$$y_{D1} = y_1 + \varepsilon_{D1} = y(\boldsymbol{x}; \boldsymbol{\theta}_1, \boldsymbol{\theta}_{12}) + \delta_1(\boldsymbol{x}) + \varepsilon_{D1} \tag{2.8}$$

$$y_{D2} = y_1 + \varepsilon_{D2} = \hat{y}_2(\boldsymbol{x}; \boldsymbol{\theta}_2, \boldsymbol{\theta}_{12}) + \delta_2(\boldsymbol{x}, y_1) + \varepsilon_{D2} \tag{2.9}$$

The Bayesian network for the quantities represented by Eqs. (2.8) and (2.9) is shown in Figure 2.2. The systematic organization of information in the Bayesian networks makes them ideal for multidisciplinary problems with several interacting models and limited data. For example, the connections between the two models are provided by inputs $\boldsymbol{x}$, parameters $\boldsymbol{\theta}_{12}$, and true output y_1 that feeds into $\hat{y}_2$. These relationships can be utilized during calibration such that the data furthest downstream (i.e., y_{D2}) can inform the uncertain parameters associated with both y_1 and y_2, including shared parameters $\boldsymbol{\theta}_{12}$. The influence of this downstream data depends on two main factors: 1) the relative amount of data available for y_{D1} and y_{D2} and 2) the dependence (i.e., sensitivity) of model $\hat{y}_2$ on y_1.

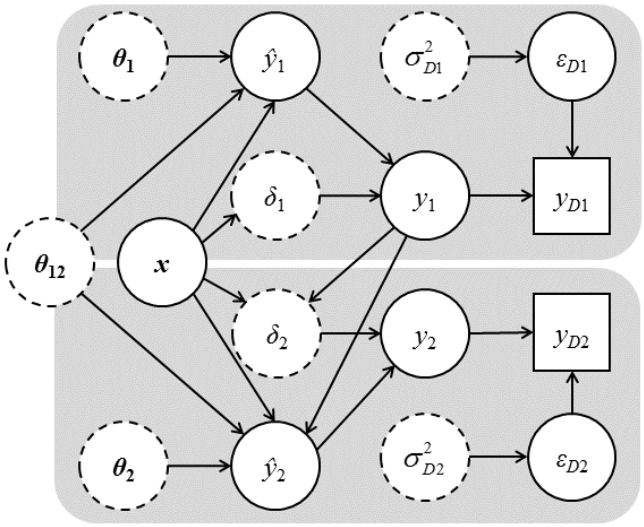

Figure 2.2. Bayesian Network for two-discipline system

Given N_1 independent observations of y_{D1} and N_2 independent observations of y_{D2} for an testing input $\boldsymbol{x}$, the simultaneous joint likelihood function of parameters $\boldsymbol{\Phi} = [\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_{12}, \delta_1(\boldsymbol{x}), \delta_2(\boldsymbol{x},y_1), \sigma_{D1}, \sigma_{D2}]$ is shown in Eq. (2.10). Assuming that the observations of y_{D1} and y_{D2} are statistically independent, Eq. (2.10) is further divided into two separable likelihoods each corresponding to a given data set.

$$\begin{aligned} L(\boldsymbol{\Phi}) &\propto \pi(y_{D1}, y_{D2} \mid \boldsymbol{\Phi}) \\ &\propto \pi(y_{D1} \mid \boldsymbol{\Phi})\pi(y_{D2} \mid \boldsymbol{\Phi}) \end{aligned} \tag{2.10}$$

The likelihood in Eq. (2.10) requires evaluating both models $\hat{y}_1$ and $\hat{y}_2$ at every input condition to generate a single posterior sample. Additionally, as discussed in Section 2.4 the number of samples required for the calibration to converge increases with the number of calibration parameters. This computational burden motivates the investigation of more efficient calibration techniques for multidisciplinary models. Therefore, in Section 2.5.1, a segmented calibration method is developed to explore its effectiveness in reducing the overall calibration effort.

2.5.1 Proposed Segmented Bayesian Model Calibration

The goal of segmented calibration is to isolate important relationships in the multidisciplinary Bayesian network, as well as make the best use of limited data and computational resources. In reference to the Bayesian network in Figure 2.2, this leads to the parameter sets $\boldsymbol{\Phi}_1 = [\boldsymbol{\theta}_1, \delta_1(\boldsymbol{x}), \sigma_{D1}]$, $\boldsymbol{\Phi}_2 = [\boldsymbol{\theta}_2, \delta_2(\boldsymbol{x},y_1), \sigma_{D2}]$, and $\boldsymbol{\Phi}_{12} = \boldsymbol{\theta}_{12}$. Modifying the Bayesian framework to calibrate in a segmented manner involves the joint likelihood function in Eq.(2.10), which is shown in its expanded form in Eq. (2.11) for $\boldsymbol{\Phi}_1$, $\boldsymbol{\Phi}_2$, and $\boldsymbol{\Phi}_{12}$.

$$\begin{aligned} L(\boldsymbol{\Phi}_1,\boldsymbol{\Phi}_2,\boldsymbol{\Phi}_{12}) &\propto \pi(y_{D1}\,|\,\boldsymbol{\Phi}_1,\boldsymbol{\Phi}_2,\boldsymbol{\Phi}_{12})\pi(y_{D2}\,|\,\boldsymbol{\Phi}_1,\boldsymbol{\Phi}_2,\boldsymbol{\Phi}_{12}) \\ &\propto \pi(y_{D1}\,|\,\boldsymbol{\Phi}_1,\boldsymbol{\Phi}_{12})\pi(y_{D2}\,|\,\boldsymbol{\Phi}_1,\boldsymbol{\Phi}_2,\boldsymbol{\Phi}_{12}) \\ &\approx \pi(y_{D1}\,|\,\boldsymbol{\Phi}_1,\boldsymbol{\Phi}_{12})\pi(y_{D2}\,|\,\boldsymbol{\Phi}_{1|D1},\boldsymbol{\Phi}_2,\boldsymbol{\Phi}_{12}) \end{aligned} \tag{2.11}$$

Since parameters $\boldsymbol{\Phi}_2$ are not connected to y_{D1} in the Bayesian network, the first likelihood in the first step of Eq. (2.11) which can exclude $\boldsymbol{\Phi}_2$ while still being equivalent to Eq. (2.10), as shown in the second step. The mechanism of segmenting the likelihood is shown between the second and last step of Eq. (2.11). The approximation sign implies that in the segmented approach, the quantity $\pi(y_{D1}|\,\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_{12})\ \pi(y_{D1}|\,\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2, \boldsymbol{\Phi}_{12})$ is approximated by $\pi(y_{D1}|\,\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_{12})\ \pi(y_{D1}|\,\boldsymbol{\Phi}_{1|D1}, \boldsymbol{\Phi}_2, \boldsymbol{\Phi}_{12})$, where $\boldsymbol{\Phi}_{1|D1}$ denotes the posteriors of $\boldsymbol{\Phi}_1$ from the first calibration. The validity of this approximation is verified in Section 2.6.1 with a global sensitivity analysis.

With a segmented likelihood in Eq. (2.11), two Bayesian posteriors emerge: one using the likelihood defined by y_{D1} in Eq. (2.12) and another using the likelihood defined by y_{D2} and conditioned on $\boldsymbol{\Phi}_{1|D1}$ in Eq. (2.13).

$$\pi(\boldsymbol{\Phi}_1,\boldsymbol{\Phi}_{12}\,|\,y_{D1}) \propto \pi(y_{D1}\,|\,\boldsymbol{\Phi}_1,\boldsymbol{\Phi}_{12})\pi(\boldsymbol{\Phi}_1,\boldsymbol{\Phi}_{12}) \tag{2.12}$$

$$\begin{aligned} \pi(\boldsymbol{\Phi}_2,\boldsymbol{\Phi}_{12}\,|\,y_{D2}) &\propto \pi(y_{D2}\,|\,\boldsymbol{\Phi}_{1|D1},\boldsymbol{\Phi}_2,\boldsymbol{\Phi}_{12})\pi(\boldsymbol{\Phi}_2,\boldsymbol{\Phi}_{12|D1}) \\ &\propto \pi(y_{D2}\,|\,\boldsymbol{\Phi}_{1|D1},\boldsymbol{\Phi}_2,\boldsymbol{\Phi}_{12})\pi(\boldsymbol{\Phi}_2)\pi(\boldsymbol{\Phi}_{12|D1}) \end{aligned} \tag{2.13}$$

The shared parameters $\mathbf{\Phi}_{12}$ are updated in both calibration segments where the posterior PDF of $\mathbf{\Phi}_{12|D1}$ from the first calibration becomes the prior of the second and assumed to be independent of $\mathbf{\Phi}_2$, as shown in Eq. (2.13). In some cases, parameters in $\mathbf{\Phi}_{12}$ can be included solely in the first or second calibration, if a sensitivity analysis determines one model is more sensitive to those parameters. A quantitative sensitivity analysis on all model outputs (see\ Chapter 4) helps determine an appropriate calibration strategy (i.e., segmented or simultaneous) for the sources of uncertainty in a multidisciplinary analysis.

2.6 Mathematical Example

In this section, segmented and simultaneous Bayesian calibration is first investigated for a mathematical example where analytical posteriors are obtained using conjugate distributions [55]. The derivations of the posterior PDFs from each calibration strategy are given in Appendix A. In Section 2.6.5, the example is extended for shared inputs and parameter θ_{12}.

First, consider two models $y_1 = \theta_1$ and $y_2 = cy_1 + \theta_2$ that are related through a dependence coefficient c. Parameters θ_1 and θ_2 are uncertain and are to be calibrated using observations y_{D1} and y_{D2} on outputs y_1 and y_2, respectively. Here, model discrepancy is assumed to be contained in the uncertainty about each model and is contained in θ_1 and θ_2 such that $y_1 = \hat{y}_1$ and $y_2 = \hat{y}_2$, as shown in Eqs. (2.14) and (2.15). The observations y_{D1} and y_{D2} contain measurement errors that are assumed to be Gaussian with zero mean with known variances σ_{D1}^2 and σ_{D2}^2, respectively.

$$y_{D1} = y_1 + N(0, \sigma_{D1}) = \theta_1 + N(0, \sigma_{D1}) \tag{2.14}$$

$$y_{D2} = y_2 + N(0, \sigma_{D2}) = cy_1 + \theta_2 + N(0, \sigma_{D2}) \tag{2.15}$$

The Bayesian network for this system is shown in Figure 2.3, where the two models are connected via nodes y_1 and y_2. The strength of their dependence is dictated by c, which influences

the viability of a segmented Bayesian calibration strategy over the network for a fixed data scenario.

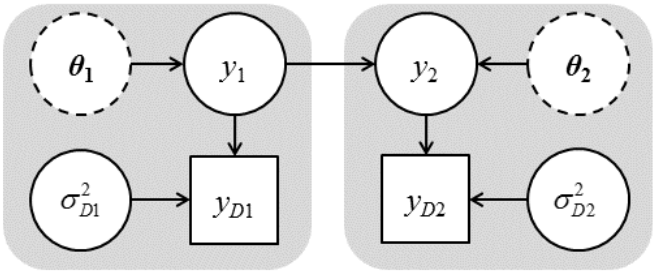

Figure 2.3. Bayesian network for mathematical example

In Section 2.6.1, prior distributions are selected for θ_1 and θ_2 and used to discuss the effect of model dependence on model output sensitivities. The posterior PDFs and output variance ratios between the segmented and simultaneous methodologies are derived in Appendix A. and the resulting distributions and downstream prediction confidence from each calibration is compared analytically in Section 2.6.2. The calibrations are then performed numerically in Section 2.6.3 using slice sampling [54] and computational effort is assessed using K-L divergence in Section 2.6.4. Finally, the demonstration on an example with shared inputs and parameters is presented in Section 2.6.5.

2.6.1 Model Dependence and Prior Sensitivities

Conjugate distributions facilitate the derivation of the segmented and simultaneous posteriors, where for simplified problems the distribution type of the posterior is the same as that of the prior when combined with an appropriate choice of the likelihood function (e.g., a normal likelihood function and normal prior result in a normal posterior) [55]. For this reason, the simple linear relationship between Eqs. (2.14) and (2.15) can be further used to derive the posterior statistics for θ_1 and θ_2 using normal priors $\pi(\theta_1) \sim N(\mu_{\theta 1}, \sigma_{\theta 1})$ and $\pi(\theta_2) \sim N(\mu_{\theta 2}, \sigma_{\theta 2})$ given in Table 2.1. In

addition, there are N_1 observations of y_{D1} and N_2 observations of y_{D2} with known observation variances σ_{D1}^2 and σ_{D2}^2 also given in Table 2.1.

Table 2.1 True parameter values, prior means and variances, and measurement error variance

Parameter	Quantity	Value
θ_1	Truth	1.2
	$\mu_{\theta 1}$	1.3
	$\sigma_{\theta 1}^2$	1.0e-2
θ_2	Truth	0.9
	$\mu_{\theta 2}$	0.7
	$\sigma_{\theta 2}^2$	1.0e-2
Measurement Error Variance	σ_{D1}^2	2.5e-3
	σ_{D2}^2	10.0e-3

The prior first order sensitivities of model y_2 to the uncertainty in θ_1 and θ_2 for dependence coefficients of c = [0, 5] are computed using Eqs. (2.16) and (2.17), and are presented in Figure 2.4. They represent the contribution to the total variance in y_2 from each source of uncertainty. Since the priors are assumed uncorrelated, their first order and total effects are equivalent. Furthermore, due to the strictly feed-forward nature of the models, the sensitivity of y_1 to θ_1 is 1 and its sensitivity to θ_2 is zero.

$$S_{1,y2}^{\theta 1} = S_{T,y2}^{\theta 1} = \frac{Var_{\theta 1}(E_{\theta 1}(y_2 \mid \theta_2))}{Var(y_2)} = \frac{c^2\sigma_{\theta 1}^2}{c^2\sigma_{\theta 1}^2 + \sigma_{\theta 2}^2} \tag{2.16}$$

$$S_{1,y2}^{\theta 2} = S_{T,y2}^{\theta 2} = \frac{Var_{\theta 2}(E_{\theta 2}(y_2 \mid \theta_1))}{Var(y_2)} = \frac{\sigma_{\theta 2}^2}{c^2\sigma_{\theta 1}^2 + \sigma_{\theta 2}^2} \tag{2.17}$$

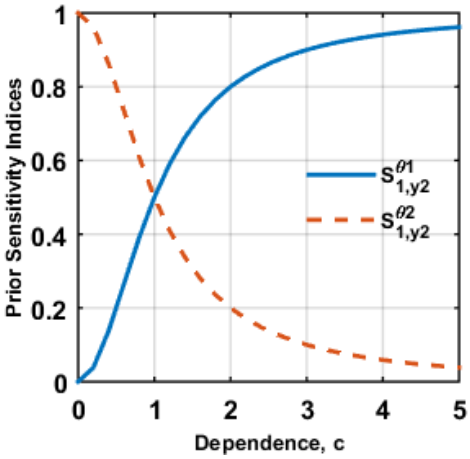

Figure 2.4. Prior sensitivity of model output y_2 to uncertainty in θ_1 and θ_2

At c = 0, parameter θ_1 has no influence on output y_2 as indicated in Figure 2.4 with zero sensitivity. Intuitively, increasing the dependence coefficient results in increased sensitivity of y_2 to the uncertainty in θ_1, and asymptotically approaches 1 as c increases. Note that at c = 1, the first order sensitivities are 0.5 for both θ_1 and θ_2 because the variance contributions from $c\theta_1$ and θ_2 on output y_2 are equal.

2.6.2 Calibration Results for Analytical Example

The posterior distributions for segmented and simultaneous calibrations are derived in Appendix A and used to test the effects of the dependence coefficient c and the relative number of data points N_1 and N_2 on the posterior parameter and prediction uncertainty. The variances of θ_1 from a segmented strategy (Eq. (A.4)) and from a simultaneous strategy (Eq. (A.22)) as a function of N_1 are shown in Figure 2.5 for two cases of downstream data: N_2 = 5 and N_2 = 50.

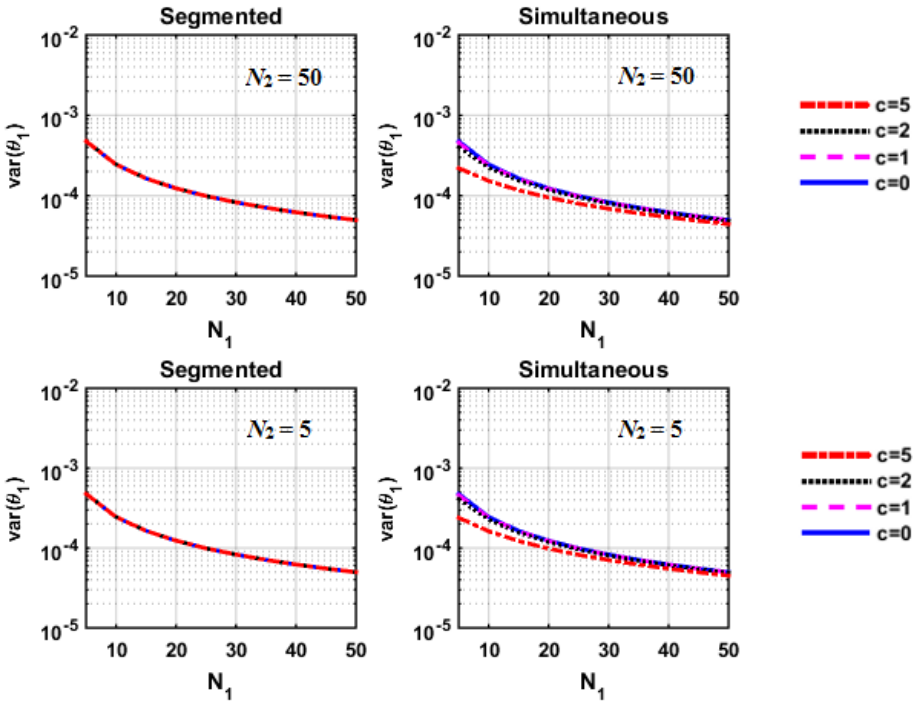

Figure 2.5. Posterior variance of θ_1 as a function of data on y_1

It is observed in Figure 2.5 that the segmented calibration of θ_1 is solely dependent on the amount of data on y_1 and is unaffected by the amount of data on y_2 or the dependence coefficient. In contrast, a simultaneous strategy does allow downstream information y_{D2} to influence θ_1 noticeably at low levels of N_1 and large c. These figures imply that as model dependence increases, downstream data becomes important and a segmented calibration strategy fails to benefit from this information. However, the effectiveness of this downstream data is limited. For example, when c = 5 the amount of uncertainty reduction in θ_1 between N_2 = 5 and N_2 = 50 for a simultaneous strategy is only 8.8%.

In contrast to the posterior variances of θ_1, c shows the posterior variances of θ_2 computed using Eqs. (A.8) and (A.16) in Appendix A. When there is ample upstream data (i.e., N_1 = 50), the posterior variance from the segmented strategy is effectively independent of c. This means that θ_1 does not influence the downstream segmented calibration when it is well-characterized. However,

when N_1 = 5, more uncertainty propagates forward leading to more uncertainty in θ_2. In a simultaneous strategy, however, the correlations between θ_1 and θ_2 posterior samples (denoted by $\rho(\theta_1, \theta_2)$), are negative and lead to observing more uncertainty in θ_2 compared to the segmented strategy. Conversely, the correlations $\rho(\theta_1, \theta_2)$ in a segmented strategy are zero.

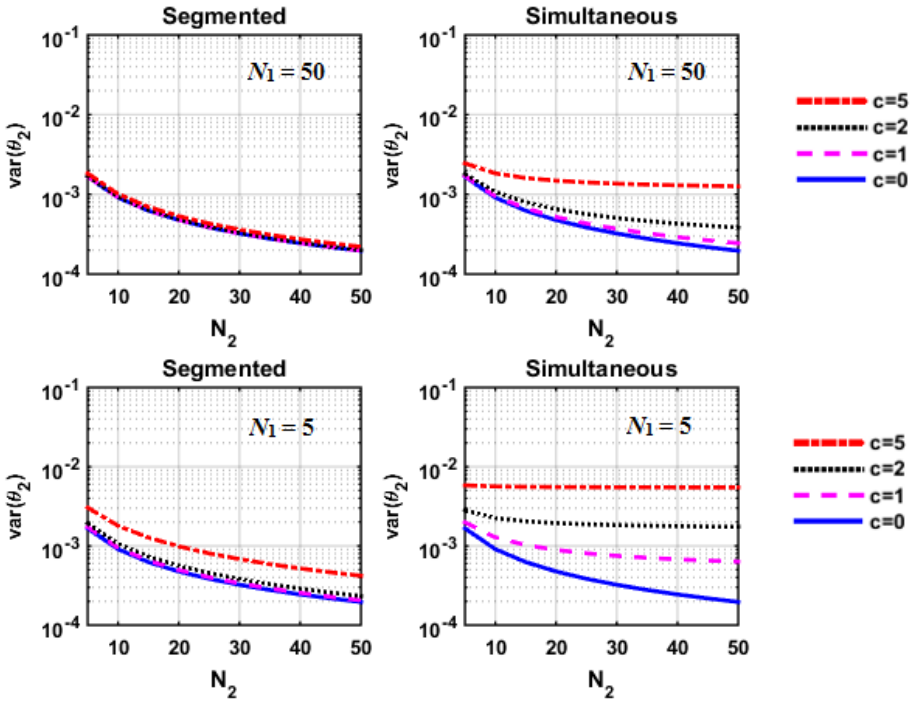

Figure 2.6. Posterior variance of θ_2 as a function of data on y_2

As shown in Figure 2.7, the simultaneous calibration correlation coefficient in Eq. (A.26) is dependent on relative data sizes corresponding to the two model outputs. The parameters exhibit strong negative correlations when N_1 is limited and with increased model dependence.

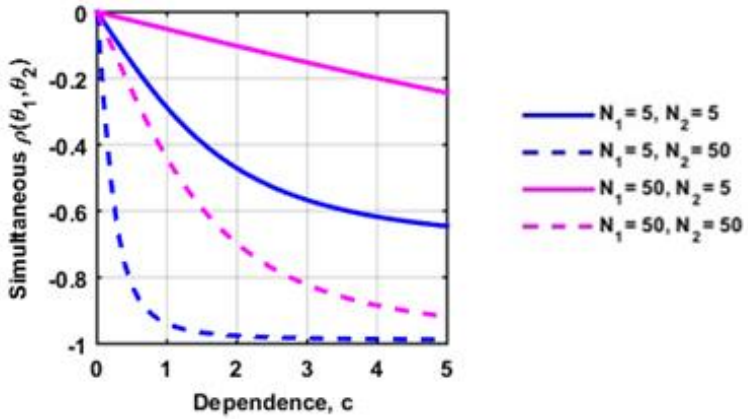

Figure 2.7. Posterior correlation between θ_1 and θ_2 as a function of *c* computed using simultaneous calibration)

To account for $\rho(\theta_1, \theta_2)$, posterior variances in y_2 for both segmented and simultaneous strategies for the analytical example are given in Eq. (2.18) and shown in Figure 2.8. Even though the variance of θ_2 from simultaneous calibration with $N_1 = 5$ is significantly greater than the variance segmented calibration of θ_2, neglecting the negative correlation between θ_1 and θ_2 in a segmented strategy results in more uncertainty propagating downstream to y_2. In other words, neglecting the correlation results in greater downstream prediction uncertainty from the segmented calibration strategy as the dependence between the models increases.

$$Var(y_2) = c^2 Var(\theta_1) + Var(\theta_2) + 2c\rho(\theta_1, \theta_2)\sigma_{\theta 1}\sigma_{\theta 2} \tag{2.18}$$

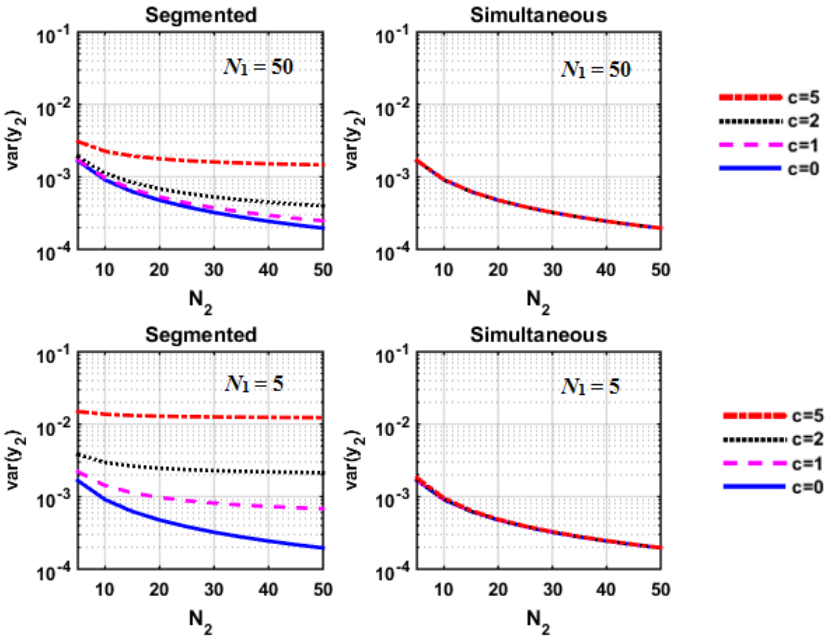

Figure 2.8. Posterior variance of y_2 as a function of data on y_2

2.6.3 Comparison of Segmented and Simultaneous Calibration Results

The posterior distributions resulting from the segmented and simultaneous calibration strategies are be compared using two metrics. Here we compare their mean values and variances; the former is related to bias and the latter to precision. The mean values from both strategies were found to be identical; thus there is no bias introduced by the segmented strategy for these examples. Figure 2.9 shows the variance ratio for the posterior distributions of θ_1, θ_2, and y_2, with the variance from the simultaneous calibration as the denominator and the variance from the segmented calibration as the numerator. A variance ratio greater than 1 indicates that the result of the segmented strategy contains more uncertainty (i.e., more imprecision) and Figure 2.9 shows these ratios for the four data scenarios mentioned earlier.

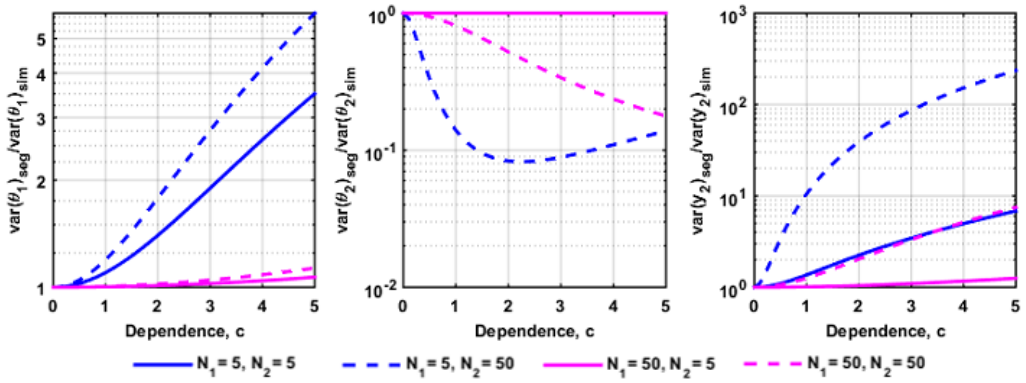

Figure 2.9. Variance ratios for θ_1, θ_2, and y_2 predictions from segmented and simultaneous calibration

The variance ratios in Figure 2.9 continue to suggest that limited upstream data impedes the performance of a segmented calibration. First, the variance ratio for $y_1 = \theta_1$ prediction (which is an increasing function of c) shows significant increase when $N_1 = 5$ (since the simultaneous strategy uses y_{D2} data to reduce the variance of θ_1). This effect of limited data on y_{D1} is worsened in downstream predictions in a segmented calibration strategy as seen from the increasing variance ratio for y_2. Regarding θ_2, the variance ratio for $N_1 = 5$ and $N_2 = 50$ is significantly less than 1 because all of the y_{D2} data is used for the posterior estimation of θ_2 (thus reducing its variance) in the segmented calibration, whereas in the simultaneous strategy the y_{D2} data is used to estimate the posteriors of *both* θ_1 and θ_2 (resulting in less variance reduction of θ_2). However, the ignoring of negative correlation between θ_1 and θ_2 means more uncertainty is propagating downstream to y_2 in a segmented calibration strategy, resulting in a larger variance ratio for y_2.

2.6.4 Comparison of Computational Effort

The Kullback-Leibler divergence integral in Eq. (2.7) was computed numerically using Monte Carlo integration with the 10^4 posterior slice samples generated from calibration. A smaller D_{KL} at a given posterior sample i means that the two posterior distributions defined by the samples up to i and up to i-1 are more similar, hence, the distribution is converging at a faster rate. It is observed that the convergence rate is not affected by the amount of data. Consider the convergence rates for

N_1 = 50 and N_2 = 50 for a dependence coefficient of 2 shown in Figure 2.10. Here convergence is assumed to be reached when log(D_{KL}) = -10.

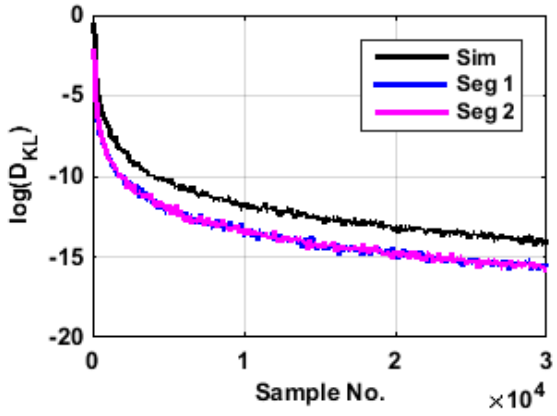

Figure 2.10. K-L Divergence for segmented and simultaneous calibration

From Figure 2.10, both segmented calibrations (the first using y_{D1} to calibrate θ_1 and the second using y_{D2} to calibrate θ_2) achieve a smaller K-L divergence than the simultaneous calibration strategy. Both segmented calibrations reached convergence in approximately 2500 samples from each model while the simultaneous converged in approximately 4500 samples from each model. This resulted in a computational savings of 2000 (45%) individual $\hat{y}_1$ simulations and 2000 (45%) $\hat{y}_2$ simulations using a segmented calibration strategy. With costly model simulations, typical of multidisciplinary analyses, these percentages can significantly influence computation time.

2.6.5 Extension: Shared Inputs and Parameters

Building on the analytical example in Eqs. (2.14) and (2.15) to more closely resemble the Bayesian network in Figure 2.2, consider the addition of shared input x and parameter θ_{12}. Thus the above mathematical example is extended further, as shown in Eqs. (2.19) and (2.20). Note that we still do not include model discrepancy terms, since those would be calibrated similarly to local parameters θ_1 and θ_2. However, model discrepancy terms are explicitly considered in Section 2.7 with aerothermal model errors.

$$y_{D1} = \theta_1 + \theta_{12} x + N(0, \sigma_{D1}) \tag{2.19}$$

$$y_{D2} = c y_1 + \theta_2 + \theta_{12} x^2 + N(0, \sigma_{D2}) \tag{2.20}$$

Data was synthetically generated for this problem using the true values of θ_1 and θ_2 and known variances σ_{D1}^2 and σ_{D2}^2 from Table 2.1 along with the true value of θ_{12} shown in Table 2.2. For the data cases considered (i.e., $N_1 = 5$, $N_1 = 50$, $N_2 = 5$, $N_2 = 50$), random realizations of y_{D1} and y_{D2} were generated at equidistant points between 0 and 1. For example, when $N_1 = 5$ one random realization of y_{D1} at each $x = [0, 0.25, 0.5, 0.75, 1]$ was used for calibration.

Table 2.2 True parameter values, prior mean, and variance of θ_{12}

Parameter	Quantity	Value
θ_{12}	Truth	0.5
	$\mu_{\theta 2}$	0.5
	$\sigma_{\theta 2}^2$	1.0e-2

Since an analytical calibration solution was not available, slice sampling was used to generate 20,000 posterior samples of θ_1, θ_{12}, and θ_2 from both segmented and simultaneous calibration strategies. The calibrated model predictions extrapolated to $x = 2$ compared to the data are shown in Figure 2.11 for the particular case when $c = 5$, $N_1 = 5$, and $N_1 = 50$. (Note that the analytical example in Sections 2.6.1-2.6.4 demonstrated that the largest sacrifice in downstream prediction accuracy from a segmented calibration strategy occurred when the dependence coefficient was large and data on the first model was low). Accuracy comparisons using Bayes factor will be made between the two calibration strategies using the known true values of y_1 and y_2 at an extrapolation point $x = 1.25$.

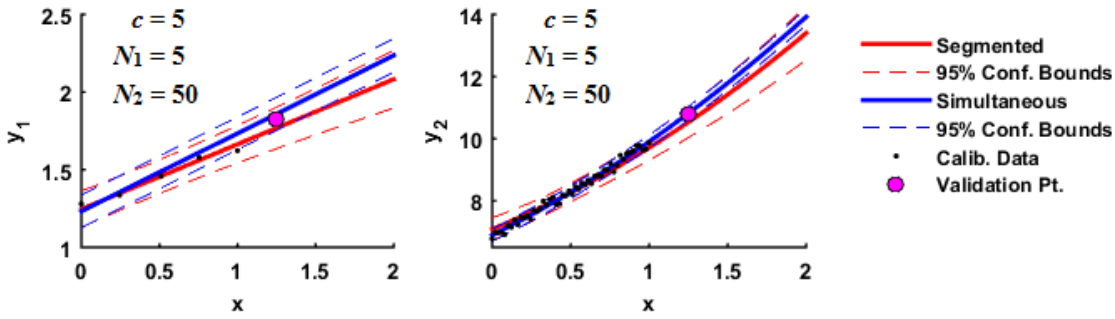

Figure 2.11. Calibrated y_1 and y_2 predictions when c = 5, N_1 = 5, and N_2 = 50

The posterior predictions in Figure 2.11 demonstrate how the extrapolation confidence in a segmented strategy is affected by the amount of upstream data available. First, the prediction of y_1 based on segmented calibration contains more uncertainty and bias than that based on simultaneous calibration. For example, at x = 1.25, variance ratio between the predictions of y_1 from the two strategies is observed to be 5.9 (i.e., segmented vs. simultaneous), and the variance ratio for y_2 is observed to be 33.3. These trends are similar to the previous example.

Figure 2.12 shows the prior and posterior distributions from the segmented and simultaneous calibration strategies compared to the true parameter values of θ_1, θ_{12}, and θ_2.The prior distribution parameters ('Prior') and true values ('True') are taken from Table 2.1 and Table 2.2.

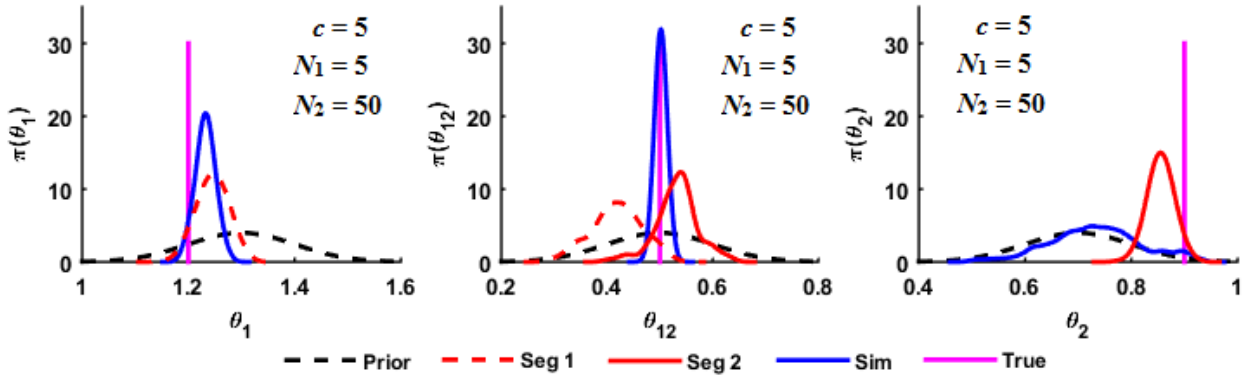

Figure 2.12. θ_1, θ_{12}, and θ_2 posterior distributions when c = 5, N_1 = 5, and N_2 = 50

The results show that the posterior distribution of θ_1 from both calibration methods contains positive bias relative to the true value of θ_1. For segmented calibration, this positive bias in θ_1 from

the first calibration ('Seg 1') is accounted for in the second calibration ('Seg 2') with negative bias in θ_2 due to their relationship in Eq. (2.20). In contrast, the posterior of θ_1 in the simultaneous calibration ('Sim') has less uncertainty and bias from the use of downstream information. Similar to the analytical example, the posterior uncertainty in θ_2 from simultaneous calibration is larger than that from segmented calibration when c = 5 due to the insensitivity of y_2 to θ_2 at higher dependence coefficients.

Second, the two posterior distributions of θ_{12} from segmented calibration (one from each segment) are both biased. The bias in θ_{12} from the first calibration segment is from limited data N_1 = 5. The second calibration of θ_{12} aims to account for the previous negative bias in y_1 data by adjusting θ_{12} positively with the information on y_2.

At x = 1.25, Figure 2.13 compares the accuracy of the calibrated models from each method for dependence coefficients ranging from 0 to 5, using Bayes factors. The Bayes factors were calculated using Eq. (2.5), where B_1 and B_2 correspond to y_1 and y_2 respectively. With the integrated likelihood from segmented calibration in the numerator, a Bayes factor less than 1 indicates that the validation data gives larger support to the simultaneous prediction.

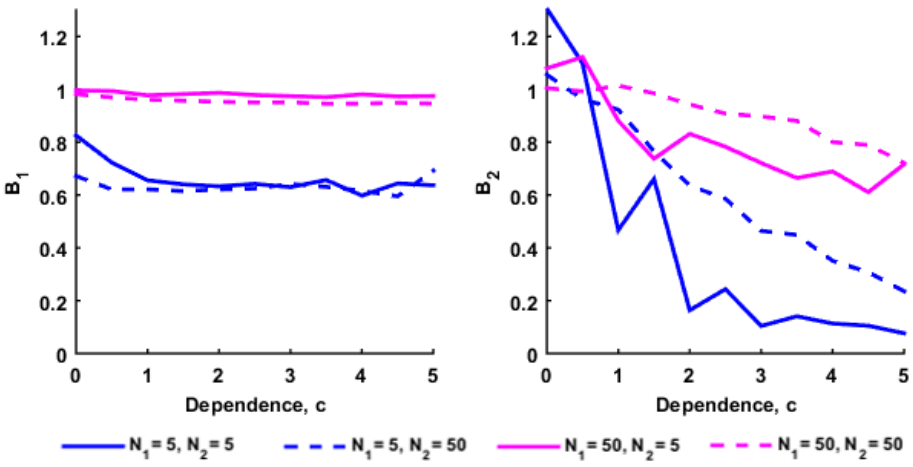

Figure 2.13. Bayes factors for y_1 and y_2 predictions at x = 1.25 from segmented and simultaneous calibration

As previously seen in the analytical example, Bayes factors in Figure 2.13 indicate that the accuracy of a segmented calibration strategy on the first model prediction is primarily a function of the data available on y_1 rather than the dependence between the models. When $N_1 = 5$, the Bayes factor B_1 across all values of c averages to 0.65 and when $N_1 = 50$ they average to 0.98. Also, the Bayes factor for y_2 is a function of both the available data and the dependence coefficient and decreases as a function of c more rapidly when there is low data on the first model ($N_1 = 5$) compared to when $N_1 = 50$.

In summary, segmented calibration is a good approximation to simultaneous calibration when the level of dependence between the models is low or when ample data is available on the first model output. Further, segmented calibration ignores negative correlation between posterior parameter sets, thus propagating more uncertainty to downstream model predictions than simultaneous calibration. However, the segmented strategy was seen to offer significant computational savings. Considering the trade-off between accuracy and computational effort, the segmented strategy appears attractive as a calibration alternative when there is ample data on the first model or when the dependence between the two models is low.

2.7 Application Problem: Aerothermal Models

The mathematical example problems presented in the preceding sections highlighted the factors that affect the calibration of multidisciplinary models, which include model dependence and relative data availability on each model output. In this section, multidisciplinary model calibration is investigated for coupled aerothermal models predicting the pressure and heating on an aircraft panel subjected to hypersonic flow conditions. The investigation focuses on whether a segmented calibration strategy is suitable for this multidisciplinary, and potentially costly, simulation.

Consider a panel on the forebody of a representative hypersonic vehicle with wedge angle θ shown in Figure 2.14. As the vehicle is subjected to hypersonic flow, an attached oblique shock with angle β is created at the leading edge (location '1'). This results in heating and aerodynamic pressure applied to the area of interest (location '4').

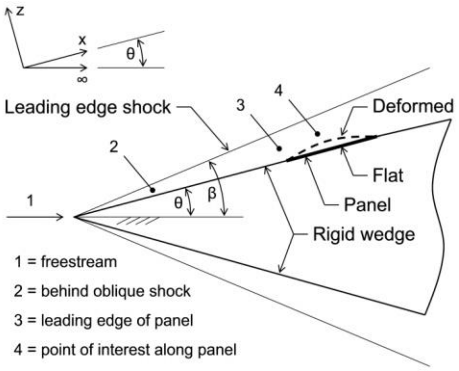

Figure 2.14. Representative hypersonic vehicle and panel deformation [28]

Figure 2.15 schematically illustrates the aerothermal interactions consisting of aerodynamic pressure and heat flux model components. Given the freestream pressure (p_1), temperature (T_1), Mach number (M_1), and panel deformation configuration (w), the local flow conditions at the panel are predicted (p_4, T_4, M_4). The aerothermal models propagate the prediction such that that the local aerodynamic pressure (p_4) is used by the aero-heating model to compute the applied heat flux (Q_4) on the panel. In a coupled aerothermoelastic analysis, upstream uncertainty and errors from model estimates propagate downstream, which impacts the accuracy of the overall aerothermoelastic response prediction. These errors further compound through time and motivate the need to identify and reduce uncertainty sources using sparse experimental data available for calibration.

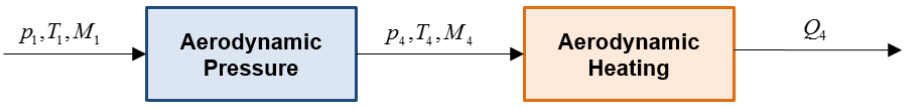

Figure 2.15. Schematic of aerothermal models

The following Sections 2.7.1-2.7.4 apply segmented and simultaneous Bayesian model calibration to an aerodynamic pressure prediction model (i.e., 3rd-order piston theory) and heating prediction model (i.e., Eckert's reference temperature method) using available high-speed wind tunnel tests. The Glass and Hunt [20] wind tunnel tests, presented in Section 2.7.2, measured both pressure and heat flux on several rigid domes with different height-to-diameter ratios intended to simulate a panel deforming into the flow. This problem provides a realistic test case for segmented calibration, since both model output quantities of interest were measured, with potential tradeoffs between computational cost and model uncertainty. A comparison of the calibration strategies is quantified with Bayes factors as well as Kullback-Leibler divergence to measure the number of samples required for posterior convergence.

2.7.1 Aerothermal Models and Bayesian network

The two models being considered in this aerothermal application problem are 3rd-order piston theory and Eckert's reference temperature method. Piston theory (*PT*) provides a simplified relationship between the unsteady pressure on the panel surface [32] which is desirable for computational tractability in aerothermoelastic predictions. The leading edge Mach number (M_3) and dynamic pressure (q_3) computed from oblique shock relations found in [56] are used in Piston theory to approximate the aerodynamic pressure load chord-wise across the panel. Piston theory accounts for both the panel slope due to deformation ($\partial w/\partial x$) and the velocity of deformation ($\partial w/\partial t$), however, since the wind tunnel specimens considered in this study are rigid (further explained in Section 2.7.2), the time-dependent terms are removed and a 3rd-order expansion of piston theory is written as Eq. (2.21).

$$p_4^{PT} = p_3 + 2\frac{q_3}{M_3}\left[\frac{\partial w}{\partial x} + \frac{\gamma+1}{4}M_3\left(\frac{\partial w}{\partial x}\right)^2 + \frac{\gamma+1}{12}M_3\left(\frac{\partial w}{\partial x}\right)^3\right] \tag{2.21}$$

After calculating the aerodynamic pressure and gas temperature T_3 (location '3' in Figure 2.14) from isentropic relations, the aerodynamic heat flux is predicted using the computationally efficient Eckert's reference temperature method assuming a calorically perfect gas [33]. Eckert's reference temperature is computed using Eq. (2.22) and the heat flux across the spherical dome is computed in Eq. (2.23).

$$T^* = T_3 + 0.5(T_w - T_e) + 0.22(T_{aw} - T_3) \tag{2.22}$$

$$Q_4^{ERT} = St^* \rho^* U_e c_p^* (T_{aw} - T_w) \tag{2.23}$$

where, St^* is the reference Stanton number, ρ^* is the reference density, U_e is the inviscid flow velocity at the dome location, c_p^* is the reference specific heat, T_{aw} and T_w are the adiabatic wall and actual wall temperature, respectively and T_e is the boundary layer edge temperature.

Consistent with the Bayesian model calibration framework presented in Section 2.2, the measurement noise in pressure ($\varepsilon_{D,p}$) and heat flux ($\varepsilon_{D,Q}$) are assumed as normal distributions. These measurement error distributions were calibrated in a previous study using the same data set assuming Gaussian distributions with zero means and variances $\sigma_{D,p}^2$ and $\sigma_{D,Q}^2$, respectively [57]. Eqs. (2.24) and (2.25) are rewritten for piston theory and Eckert's reference temperature.

$$p_D = p_4 + \varepsilon_{D,p} = p_4^{PT} + \delta_p^{PT} + \varepsilon_{D,p} \tag{2.24}$$

$$Q_D = Q_4 + \varepsilon_{D,Q} = Q_4^{ERT} + \delta_Q^{ERT} + \varepsilon_{D,Q} \tag{2.25}$$

Figure 2.16 shows the Bayesian network corresponding to the relationships between the aerodynamic pressure and heat flux model predictions along the panel (p_4, Q_4), aerothermal data (p_D, Q_D), model inputs (p_1, M_1), random model inputs (T_1, T_w), measurement errors ($\varepsilon_{D,p}$, $\varepsilon_{D,Q}$), and discrepancy terms for calibration (δ_p^{PT}, δ_Q^{ERT}).

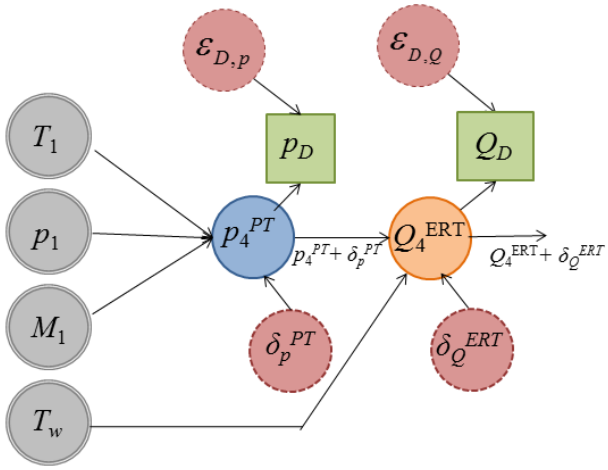

Figure 2.16. Bayesian network for aerothermal models and experiments

The discrepancy models are chosen as a function of dome slope following previous work [57] (see Section 2.1) and given in Eqs. (2.26) and (2.27). The prior uniform distribution parameters for the uncertain model discrepancy coefficients are given in Table 2.3.

$$\delta_p^{PT} = b_0^{\mathrm{PT}} + b_1^{\mathrm{PT}} \frac{\partial w}{\partial x} + b_2^{\mathrm{PT}} \left(\frac{\partial w}{\partial x}\right)^2 \tag{2.26}$$

$$\delta_Q^{ERT} = c_0^{\mathrm{ERT}} + c_1^{\mathrm{ERT}} \frac{\partial w}{\partial x} + c_2^{\mathrm{ERT}} \left(\frac{\partial w}{\partial x}\right)^2 \tag{2.27}$$

Table 2.3. Prior distributions for aerothermal error parameters

Parameter	Lower Bound	Upper Bound
b_0^{PT} (Pa)	-4e2	4e2
b_1^{PT} (Pa)	-1e4	1e4
b_2^{PT} (Pa)	-1e5	1e5
c_0^{ERT} (W/m^2)	-2e4	2e4
c_1^{ERT} (W/m^2)	-5e5	5e5
c_2^{ERT} (W/m^2)	-5e6	5e6

A global sensitivity analysis using these prior model discrepancy coefficients was performed in [57] and showed that the sensitivity of Q_4 to δ_p^{PT} and δ_Q^{ERT} are 0.592 and 0.333, respectively. These sensitivities correspond to a dependence coefficient of 1.3 in reference to Figure 2.4 indicating that a segmented calibration strategy could be viable without significantly compromising calibration accuracy. (This is of course very rough preliminary reasoning since the analytical example is not the same as the current problem). In a segmented calibration procedure the aerothermal calibration parameters are subdivided into pressure and heat flux calibration sets $\mathbf{\Phi}_p = [b_0^{PT}, b_1^{PT}, b_2^{PT}]$ and $\mathbf{\Phi}_Q = [c_0^{ERT}, c_1^{ERT}, c_2^{ERT}]$. The next subsection will describe the high-speed wind tunnel tests and aerothermal data that will be used for model calibration.

2.7.2 Aerothermal Wind Tunnel Data

Tests conducted by Glass and Hunt in 1986 at NASA's 8ft High-Temperature Wind Tunnel (HTT) investigated the thermal and structural loads on body panels in hypersonic environments [20]. These tests measured the aerodynamic pressure and heating on spherical dome protuberances that simulated deformed aircraft panels. The flow conditions for the tests of interest had a turbulent boundary layer at the panel location, and the panel holder had a sharp leading edge, similar to the representative hypersonic vehicle depicted in Figure 2.14. The spherical domes were constructed at three different height-to-diameter (*H/D*) ratios with a dome diameter of 14 in (0.356 m). The freestream hypersonic flow conditions and dimensions of the three spherical dome configurations considered in this analysis are presented in Table 2.4.

Table 2.4. Experimental conditions for three tests by Glass and Hunt [20] with different *H*/*D* ratios

Test	p_1 (Pa)	M_1	H/D
Run 30	655.0	6.60	0.028
Run 31	648.0	6.60	0.013
Run 32	655.0	6.60	0.006

Along with the Mach number (M_1) and freestream pressure (p_1) for each run, the data reports both the aerodynamic pressure (p_D) and aerodynamic heat flux (Q_D) at 58 instrumented locations on the spherical dome. For the purposes of this analysis, only 11 points along the dome centerline are considered. Therefore, there are $N_p = N_Q = 33$ data points total (i.e., 11 points on 3 domes) of aerodynamic pressure and heat flux.

Note, however, that the freestream temperature (T_1) – a shared model input – and the initial wall temperature (T_w) were left unreported in the Glass and Hunt data. The lack of information on these two parameters presents epistemic input uncertainty; however sensitivity analyses from related studies by DeCarlo et. al [57] and Smarslok et. al [58] have determined that a 10% coefficient of variation in these parameters was not a significant source of uncertainty in pressure or heat flux. As such, they will not be considered for calibration; thus, only the model discrepancy parameters identified in Section 2.7.1 are updated.

The data are subdivided into calibration and validation sets. In this objective, calibrations are performed using data from two domes with lower *H/D* ratios (i.e., Runs 31 and 32) and then the predictions based on the results of the two calibration strategies are compared against the pressure and heat flux measurements from Run 30 which had the highest *H/D* ratio. Further, a feature of Run 31 is exploited to test the effects of data availability on calibration. The front, back, and middle points of the Run 31 test set represent the most positive, most negative, and zero slopes. These

three points are used to illustrate the sparse data case; in contrast, the abundant data case has 22 data values (i.e., at 11 locations on each dome). The four calibration cases considered are shown in Table 2.5.

Table 2.5. Calibration cases with varying data available on p_D and Q_D

Case	N_p	N_Q
1	3	3
2	3	22
3	22	3
4	22	22

2.7.3 Simultaneous and segmented calibration results

Simultaneous and segmented calibrations for error parameters δ_p^{PT} and δ_p^{ERT} were conducted. For simultaneous calibration, 50,000 samples of the posterior distributions were generated using slice sampling [54]. Similarly, for segmented calibration 50,000 samples were generated for each segment.

As a representative parameter for comparing the results of the segmented and simultaneous calibration, the posterior distributions of the piston theory error parameter b_0^{PT} with $N_p = 3$ and 22 is shown in Figure 2.17. As was the case in the analytical example in Section 2.6, the posterior distributions from the first calibration are not affected by the calibration type or the model dependence but by the amount of data available. With a low dependence coefficient c, the amount of downstream data N_Q does not impact the posterior PDFs of b_0^{PT}.

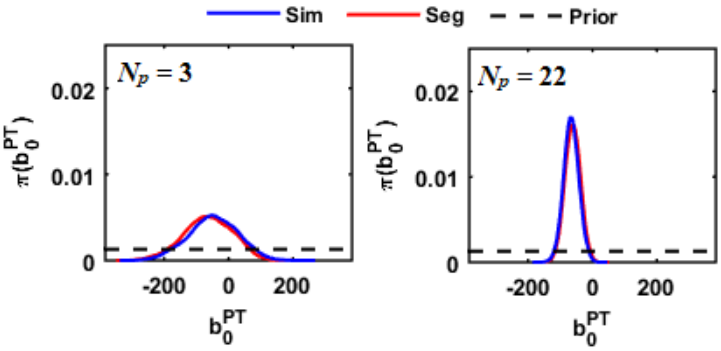

Figure 2.17. Prior and posterior distributions for b_0^{PT} from segmented and simultaneous calibration

In a segmented strategy, the posterior distributions of b_0^{PT}, b_1^{PT}, and b_2^{PT} propagate forward in the next calibration and do not change; the effect of this is seen in the downstream calibration of c_0^{ERT} shown in Figure 2.17. For the aerothermal example, when there is limited data for the first calibration but abundant data for the second calibration ($N_p = 3$, $N_Q = 22$), the segmented calibration strategy does not capture c_0^{ERT} as precisely as the simultaneous strategy. This parallels the conclusions of the analytical example when considering low model dependence (Figure 2.12).

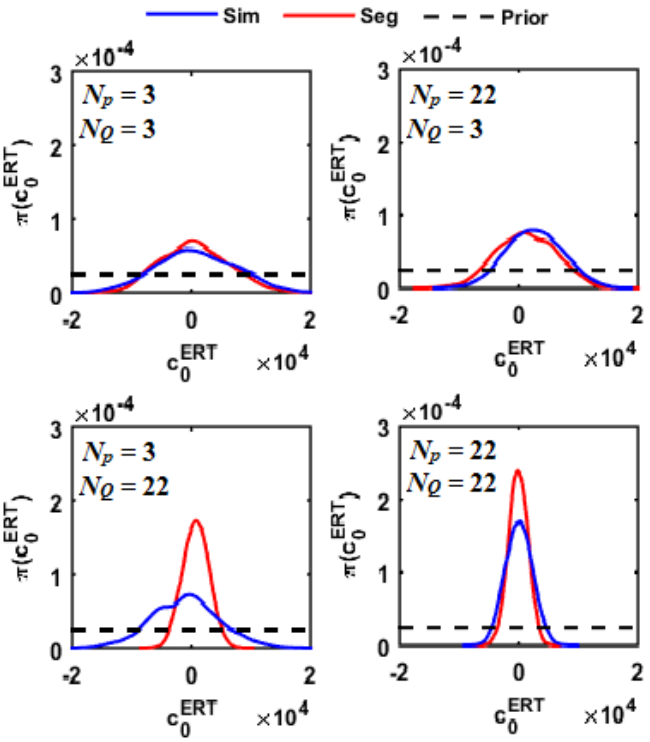

Figure 2.18. Prior and posterior distributions for c_0^{ERT} from segmented and simultaneous calibration

The posterior distributions of $\mathbf{\Phi}_p$ and $\mathbf{\Phi}_Q$ are propagated through piston theory and Eckert's reference temperature models and compared against validation pressure and heat flux data from Run 30 in Figure 2.19 and Figure 2.20, respectively. In both figures, the prediction variance is smaller at the dome center and increases towards the edges; this is because of the model discrepancy form in Eqs. (2.26) and (2.27). The slope is zero at the dome center, thus only the first terms in Eqs. (2.26) and (2.27) contribute to the prediction variance; as we move towards the edges, the slopes are non-zero and increasing, thus increasing the prediction variance.

Similar to the posterior results for b_0^{PT} in Figure 2.17, the posterior pressure prediction is affected by the amount of data N_p; the prediction bounds are wider for the sparse data case. The prediction bounds from both calibration strategies are generally similar; the slight differences

towards the edges are explained by the differences in the mean predictions and the correlation coefficients among the calibration parameters estimated by the two strategies.

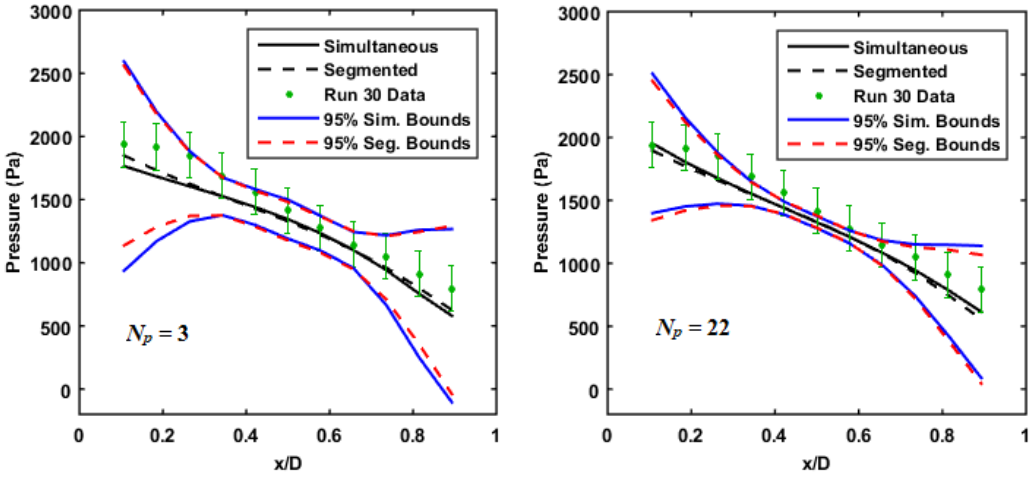

Figure 2.19. Pressure prediction vs. observation across dome when N_p = 3 and 22

However, uncertainty in the downstream heat flux prediction increases from the segmented calibration. For example, when N_p = 3 in Figure 2.19 the variance at the dome edges from segmented calibration is significantly greater than those from the simultaneous calibration. When there is ample N_p data, however, the segmented and simultaneous strategies have small differences. These differences are quantified in Section 2.7.4. using the Bayes factor.

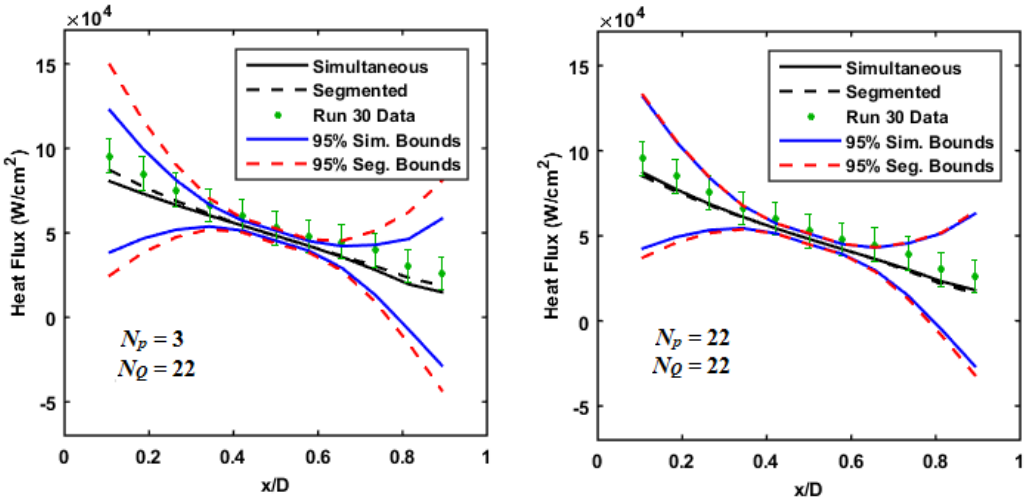

Figure 2.20. Eckert's reference temperature predictions across dome from posterior distributions when N_Q = 22

The explanation for the differences between Eckert's reference temperature predictions in Figure 2.20 is that segmented calibration ignores negative correlation between $\mathbf{\Phi}_p$ and $\mathbf{\Phi}_Q$ as discussed in Section 2.7.4. These correlations are significant in the simultaneous posterior samples and are dependent on the relative amount of observed data; for the data cases considered, the correlation coefficient between parameters b_0^{PT} and c_0^{ERT} are listed in Table 2.6. Not accounting for these correlations in a segmented strategy leads to additional variance in downstream predictions, similar to the analytical example in Section 2.6.

Table 2.6. Correlation coefficients between b_0^{PT} and c_0^{ERT} from simultaneous calibration

N_p	N_Q	ρ
3	3	-0.74
3	22	-0.97
33	3	-0.28
33	22	-0.74

2.7.4 Comparison of Calibration Strategies

To assess the effectiveness of the segmented calibration methodology compared to the simultaneous procedure, this subsection compares the accuracy and computational effort of the segmented and simultaneous calibration methods. Accuracy is compared using Bayes factors, and computational effort is compared using convergence of the K-L divergence metric.

Values of Bayes factors *B* are computed at each Run 30 dome location using Eq. (2.5). Table 2.7 shows average *B* values over all locations, *B* value at the dome center, and *B* value at the front edge of the dome. In the last column, the Bayes factors for the pressure and heat flux prediction (B_p and B_Q) *at each location* are multiplied and averaged over all locations to compute $B_{total,ave}$. A Bayes factor greater than 1 indicates that the data favors the segmented calibration method.

Table 2.7. Bayes factors across the dome for pressure and heat flux predictions

N_p	N_Q	$B_{p,ave}$	$B_{p,mid}$	$B_{p,front}$	$B_{Q,ave}$	$B_{Q,mid}$	$B_{Q,,front}$	$B_{combined,ave}$
3	3	1.06	0.92	1.22	1.04	0.83	1.02	1.11
3	22	1.07	0.85	1.17	1.02	0.95	0.84	1.08
22	3	0.94	1.04	0.99	0.99	0.90	0.96	0.93
22	22	0.95	1.02	0.97	0.94	1.01	0.91	0.90

Most of the Bayes factors are close to 1, implying that a segmented calibration strategy results in no significant loss of accuracy for the aeropressure and aeroheating relationship. Substantial difference between competing models is only indicated by Bayes factors greater than 3, as pointed out by Jeffreys in [59].

Calibration convergence rates are computed using K-L divergence, presented in Eq. (2.7), with the integral being evaluated numerically using Monte Carlo integration with the slice samples generated from calibration, as explained in Section 2.4. For the aerothermal calibration, convergence was tested after every 1,000 slice samples and assumed to be reached when $\log(D_{KL,i}) \leq -8$. It is seen that the convergence rate is not affected by the amount of observed data. The average number of samples (considering four different data availability cases) at which the simultaneous and segmented procedures reached convergence are presented in Table 2.8.

Table 2.8. Average number of samples to convergence

Procedure	Model	Samples
Simultaneous	Both Models	39,000
Segmented	Piston Theory	10,000
	Eckert's Ref. Temp.	10,000

Table 2.8 shows that to obtain convergence in simultaneous calibration, both models must be evaluated 39,000 times. Whereas, for segmented calibration to obtain convergence, only 10,000 evaluations of each model were required, thus indicating substantial savings in computational effort.

2.8 Conclusion

A segmented Bayesian model calibration approach for multidisciplinary models was investigated as an alternative to full, simultaneous calibration in order to reduce the computational cost of calibration. The study included identifying the required characteristics of the data and coupled simulation (i.e. low model dependence and ample data specifically on the first model output), identifying the appropriate uncertain parameters and errors for calibration with the segmented process, and assessing the efficiency and accuracy of segmented model calibration.

The aerothermal problem was segmented into piston theory and Eckert's reference temperature model components. To quantify the viability and potential benefit of isolating calibrations of models in the Bayesian network, segmented and simultaneous calibration were compared using the Kullback-Leibler divergence and Bayes factor metrics. The Kullback-Leibler divergence was used to monitor calibration convergence, and the Bayes factor was used to assess the accuracy.

The following insights were obtained based on comparison between the two calibration approaches, using the analytical example and the aerothermal application problem.

1. As the coupling strength between the two models increases, the segmented approach loses accuracy for both upstream and downstream predictions.
2. When there is limited data available on the first model output, more uncertainty propagates downstream from the calibration of the first model parameters which affects downstream prediction confidence. Furthermore, a simultaneous strategy has the opportunity to use downstream data to reduce the uncertainty in upstream parameters. This effect becomes apparent in limited data cases and as dependence between models increases.

3. Parameter correlations inherent in the simultaneous calibration are ignored in the segmented approach which yields greater downstream prediction uncertainty.
4. For problems where coupling strength is not readily obvious (e.g., the aerothermal example), a global sensitivity analysis can indicate the coupling strength.
5. When the two models have shared parameters, the shared parameters are calibrated twice in a segmented calibration strategy (i.e., the posterior from the first calibration is the prior for the second calibration). This offers a second opportunity for uncertainty reduction in the shared parameter.

The calibration convergence comparison using K-L divergence indicates higher computational efficiency of the segmented approach. For the application problem, the reduction in the number of evaluations of each model using the segmented approach compared to simultaneous calibration was 74.4%. Further, in terms of accuracy comparison, both strategies yielded similar posterior predictions, as indicated by the Bayes factors.

Further investigation into the tradeoffs between accuracy and efficiency between the segmented and simultaneous calibration methods is needed, especially when different models require different computational effort. Additional work on multidisciplinary model calibration needs to address transient simulations and feedback coupling between disciplinary models. In the presence of feedback coupling, sensitivity analyses and forward propagation of uncertainty are not straightforward. Future work may also address resource allocation for data collection for multidisciplinary model calibration, and its connection to the preferred calibration strategy.

CHAPTER 3

MODEL DISCREPANCY CALIBRATION IN TIME-DEPENDENT, COUPLED ANALYSES

3.1 Introduction

This chapter is motivated by the challenge of isolating model discrepancy in coupled, time-dependent analyses where model errors accumulate across the coupling interfaces and through time. Quantifying single-disciplinary model error contributions in multidisciplinary analyses (MDA) was addressed in Chapter 2, however, that chapter focused on choosing a Bayesian model calibration strategy (i.e., segmented or simultaneous calibration) when multiple sources of calibration data are available. The segmented calibration strategy will be effective only if data is available for each calibration segment. However, data is often only observed on one output QoI (i.e., wall temperature), which is often seen in multidisciplinary applications where one data source is used to calibrate potentially numerous sources of uncertainty [60,61]. Therefore, this chapter investigates effective error parameterization and aggregation strategies to isolate error contributions from coupled models through time with data limited to one QoI.

Central to calibrating aeropressure and aeroheating models in Chapter 2, however, was determining how to best represent the model discrepancy that propagates through the Bayesian network. Often in Bayesian applications, model discrepancy is formulated to mitigate parameter bias during calibration of an unobservable QoI (e.g., structural damping). For this purpose, Kennedy and O'Hagan (KOH) developed a model calibration framework that used Gaussian process (GP) models to account for model inadequacy [62]. The GP representation of model discrepancy in KOH may also include a trend function term, which can be used to parametrically represent the dependence of discrepancy in model inputs.

Related calibration efforts by Smarslok et al [58] parameterized the static aeropressure and aeroheating model discrepancies as systematic biases which improved the agreement between the coupled prediction and data the at the dome centers. However, these error formulations were not sufficient to capture the prediction error among the remaining dome centerline points (see Section 2.7.2). After observing the dependence of the prediction error on the dome slope, model discrepancies for both piston theory and Eckert's reference temperature method were then represented as functions of slope and inferred using a static Bayesian network.

When data through time is available on one QoI and the entire transient analysis is treated as a black box, the cumulative effects of the propagating errors are viewed globally across all time instants. However, this global representation of the model discrepancy does not identify significant sources of model error within the coupled analysis. Instead, model errors that are observed at each time instant can be corrected in a step-wise fashion, however, this approach would treat the coupled prediction as a black box and is unable to isolate individual model discrepancy contributions. What is needed is a partitioned approach to model discrepancy, where each model output at each time instant in the analysis is corrected before propagating to the next model. Again, one significant challenge is that data is often available only on one output QoI. Thus, this chapter investigates the use of dynamic Bayesian networks to capture individual model discrepancy contributions in time-dependent, multidisciplinary analyses.

This chapter applies these ideas to an aerothermal analysis where errors aggregate between Eckert's reference temperature method and heat transfer over a dynamic Bayesian network in Section 3.2. Time-dependent temperature data from historic hypersonic wind tunnel experiments is presented in Section 3.3 and applied to the dynamic Bayesian network for calibration. Three different model discrepancy resolutions are developed in Section 3.4 (global, step-wise, and

partitioned approaches), calibrated using the time-dependent temperature data over the dynamic aerothermal Bayesian network in Section 3.5, and compared using the model reliability metric to assess post-calibration prediction confidence.

3.2 Partitioned Aerothermal Models

The relationships between aerodynamic pressure, aerodynamic heating, and heat transfer are shown in Figure 3.1, where there is a feed-back relationship between heat flux Q_4 and the wall temperature T_w. Specifically, this chapter focuses the analysis on the aerodynamic heating and heat transfer model components in Figure 3.1 which is a quasi-static, coupled analysis shown in Figure 3.2 in further detail.

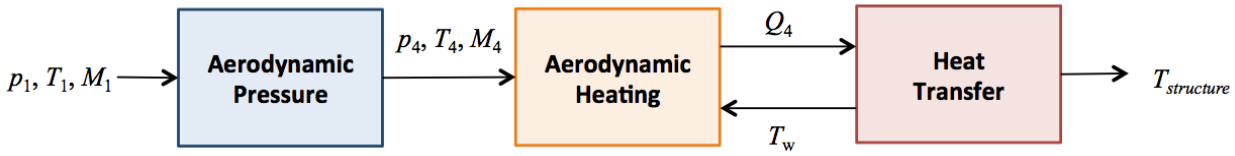

Figure 3.1. Time-Dependent Aerothermal Coupling

The quasi-static aerothermal analysis steps iteratively between heat flux Q_n and wall temperature T_n predictions through n aerothermal time-steps at a rate of Δt_{AT}. In the previous chapter, aerodynamic pressure model discrepancy was calibrated as a function of the deformed slope of the panel dw/dx using Glass and Hunt pressure data p_x^{obs} and a 3rd order expansion of piston theory [32]. Here, the mean posterior aerodynamic pressure prediction p_x at location x from the segmented calibration strategy (see Section 2.7.3) is used with the initial wall temperature T_0 as an input the partitioned aerothermal analysis in Figure 3.2.

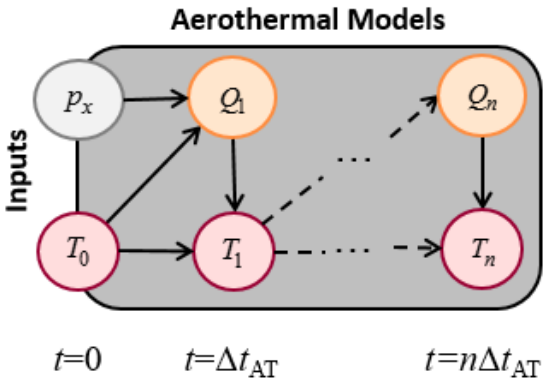

Figure 3.2. Partitioned aerothermal analysis

The model components considered in Figure 3.2 are Eckert's reference temperature method for predicting heat flux Q_n and one-dimensional heat transfer to predict wall temperature T_n. Assuming a calorically perfect gas [63], Eckert's reference temperature T^* at location x and time-step n is computed using Eq. (3.1) where T_n^{aw} and T_n represent the adiabatic wall and current wall temperature, respectively, and T_n^{bl} and T_n^{e} represent the boundary layer edge temperature at the leading edge of the panel (x = 0). Eckert's reference temperature method predicts the applied heat flux Q_{n+1}at time-step n+1 using the reference Stanton number St_n^*, the reference flow density ρ_n^*, the reference specific heat $C_{p,n}^*$ along with the inviscid flow velocity U_n as shown in Eq. (3.2).

$$T_n^* = T_n^e + 0.5(T_n - T_n^{bl}) + 0.22(T_n^{aw} - T_n^e) \tag{3.1}$$

$$Q_{n+1} = St_n^* \rho_n^* U_n C_{p,n}^* (T_n^{aw} - T_n) \tag{3.2}$$

The wall temperature T_{n+1} at the $(n+1)^{th}$ time-step is predicted with one-dimensional heat transfer using the current heat flux Q_{n+1} and wall temperature T_n from the previous time step. One-dimensional heat transfer is applicable to the thin-walled spherical dome protuberances and is shown in Eq. (3.3). Assuming a uniform initial wall temperature T_0 = 300K at t = 0, Eqs. (3.1) through (3.3). are solved explicitly at each aerothermal time step of Δt_{AT} =0.05s according to the sequence shown in Figure 3.2.

$$T_{n+1} = T_n + \frac{Q_{n+1}(T_n)\Delta t_{AT}}{\rho_{dome} C_{p,dome} \tau_{dome}} \tag{3.3}$$

Next, time-dependent temperature data is generated from three fully-turbulent aerothermal wind-tunnel tests in Section 3.3. Then, three model discrepancy resolutions are developed in Section 3.4 to capture for model error aggregation through time. Time-dependent temperature data is integrated into a dynamic Bayesian network in Section 3.5 to calibrate model discrepancy for three model discrepancy resolutions are compared based on prediction performance.

3.3 Time-Dependent Temperature Data

The Glass and Hunt HTT calibration experiments used in Chapter 2 is are used in this chapter to synthetically generate time-dependent temperature data, and are in part presented again for convenience. Three tests (Runs 30, 31, and 32) subjected 14-in diameter rigid domes to a fully turbulent boundary layer (TBL) at a surface inclination angle $\theta = 5°$. The freestream Mach number M_1, pressure p_1, total temperature $T_\infty{}^t$, and height-to-diameter ratios *H/D* for these tests are presented in Table 3.1. With the highest *H/D* ratio, the Run 30 dome represented extreme panel deformation into the flow while Run 32 with the lowest *H/D* ratio represented shallow deformation.

Table 3.1. Experimental conditions from Glass and Hunt for TBL tests with different *H*/*D* ratios

Test	M_1	p_1, Pa	$T_\infty{}^t$,°R	H/D
Run 30	6.60	655.0	3590	0.028
Run 31	6.60	648.1	3460	0.014
Run 32	6.60	655.0	3530	0.007

The domes were instrumented with pressure sensors and thermocouples at 58 locations and subjected to approximately 5 seconds of freestream Mach 6.60 flow. The 11 centerline surface pressure measurements p_x and temperature measurements T_x (x = 1 to 11) parallel to the flow are used for calibration and validation. Measurements were taken at a rate of 20 samples per second, however, Glass and Hunt reported the resulting heat loads Q_x^{obs} by transforming the temperature history using the linear heat transfer relationship $Q_x^{obs}=\rho C_p \tau\, \Delta T_x/\Delta t_{obs}$. Here, τ_{dome} is the dome thickness of 0.00157m, ρ_{dome} and $C_{p,dome}$ are the density and specific heat of aluminum (7000 series), and Δt_{obs} is the temperature measurement rate (i.e., Δt_{obs} = 0.05s).

Since the objective of this chapter is to isolate model discrepancy in transient, aerothermal predictions, a five second temperature history $T_{x,n}^{obs}$ (n = 1 to 100) was reconstructed from the reported heat flux Q_x^{obs} using Eq. (3.4). An initial wall temperature $T_{x,0}$ of 300K was assumed uniform along the dome centerlines and measurement error ε_{obs} was added as a Gaussian random variable with zero mean and a standard deviation σ_{obs} of 1K to generate time-dependent temperature data.

$$T_{x,n}^{obs} = T_{x,n-1} + \frac{Q_x^{obs}\Delta t_{obs}}{\rho_{dome} C_{p,dome} \tau_{dome}} + c_x t^2 + \varepsilon_{obs} \tag{3.4}$$

An additive quadratic term $c_x t^2$ in Eq. (3.4) introduced synthetic error into the reconstructed temperature history. The coefficient c_x was chosen such that the temperature data reached thermal equilibrium ($\partial T/\partial t$= 0) at the end of the 5 second time-history (c_x=-Q_x^{obs}/10$\rho C_p \tau$ Ks^{-2}). The temperature profiles at the front, midpoint, and back of the Run 30, 31, and 32 domes are shown in Figure 3.3. The large heat fluxes reported at the front of the domes intuitively led to elevated temperatures at these locations. Also, the dome midpoints (when dw/dx = 0) for all three runs have similar temperature profiles since Q_x^{obs} at those locations were of similar magnitude.

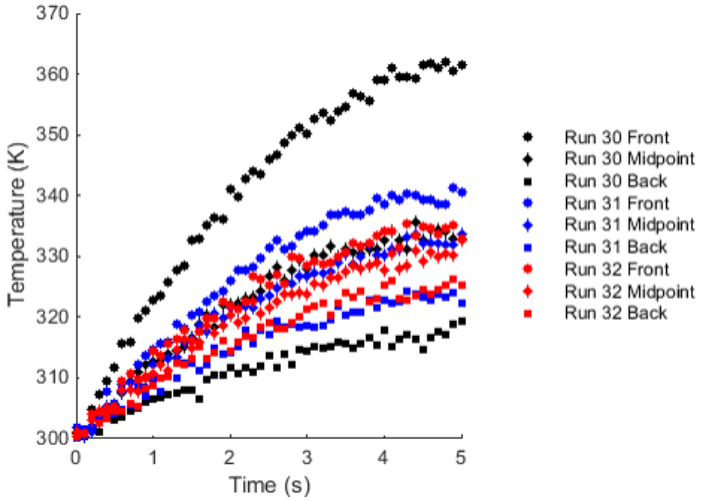

Figure 3.3. Synthetically generated temperature history for Runs 30, 31, and 32 from t = 0 to 5 seconds.

Furthermore, Figure 3.3 indicates three general phases of temperature evolution from t = 0 to 5s in the reconstructed temperature histories: 1) near-linear from 0 to 1 seconds when the effective heat flux $Q_{x,n}^{eff}$ remains approximately equal to Q_x^{obs} 2) non-linear from approximately 1 to 4 seconds as the effective heat flux $Q_{x,n}^{eff}$ decreases from the reported Q_x^{obs} 3) near-constant from 4 to 5 seconds while effective heat flux $Q_{x,n}^{eff}$ approaches 0 at t = 5s. For this reason, the temperature histories were subdivided into calibration and validation sets where the temperatures observed from 1 to 4 seconds were used for calibration and the remaining two seconds of data (from 0 to 1 second and 4 to 5 seconds) are used for comparing the posterior predictions. Other implications of this division on error parameterization and calibration are discussed in Section 3.5.

Since the temperature history at front and back of the Run 30 dome bounds the temperature profiles of the other two domes in Figure 3.3, the temperature history at these locations is compared against the partitioned aerothermal model predictions in Figure 3.4.

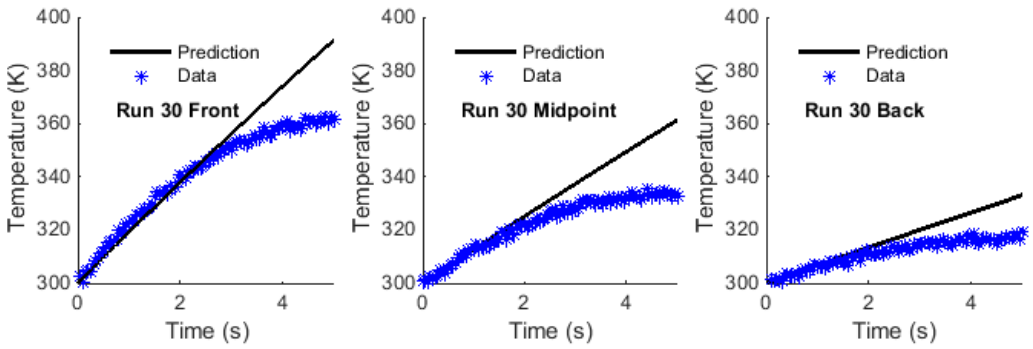

Figure 3.4. Nominal aerothermal temperature predictions and data at Run 30 dome front, midpoint, and back

The plots in Figure 3.4 highlight that the discrepancy between the temperature data T^{obs} and the temperature prediction T grows through time as the data reaches the imposed equilibrium. Section 3.4 identifies three model discrepancy resolution options for error propagation and inference and Section 3.5 constructs a dynamic Bayesian network to calibrate the prediction inadequacies in Figure 3.4.

3.4 Model Discrepancy Resolutions for Transient, Coupled Analysis

Model discrepancy δ is present in both heat flux and temperature predictions - δ_Q from Eckert's reference temperature in Eqs. (3.1) and (3.2) and δ_T from 1D heat transfer in Eq. (3.2) – which aggregate across the coupling interfaces and through time. The current approach focuses on identifying the best parameterization and resolution of coupled model discrepancy for calibration when data is limited to one output QoI.

Sections 3.4.1 through 3.4.3 present the global, step-wise, and partitioned model discrepancy resolutions for transient, coupled analysis using the aerothermal application from Section 3.2. These three discrepancy resolutions are investigated for their ability to capture the aerothermal model discrepancies δ_Q and δ_T that propagate through the dynamic Bayesian network (DBN) constructed in Section 3.5 and their prediction performance in an extrapolation region.

3.4.1 Global Model Discrepancy

A global model discrepancy strategy is shown in Figure 3.5 where the coupled model prediction is treated as a black box throughout the entire simulation time history from $t = 0$ to $n\Delta t_{AT}$. After n time-steps, the aerothermal wall temperature prediction $T_{w,n}$ is corrected to $T_{w,n}^{*}$ using an additive global discrepancy function δ as shown in Eq. (3.5).

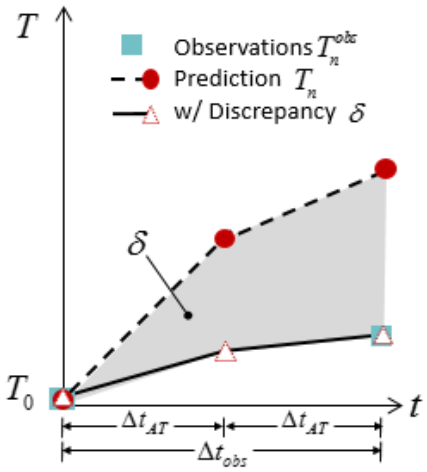

Figure 3.5. Global model discrepancy through time

$$T_{w,n}^{*} = T_{w,n}(T_{w,n-1}, Q_t) + \delta(T_{w,n-1}, Q_n) \tag{3.5}$$

Note that the global discrepancy δ in Eq. (3.5) does not delineate error contributions from heat flux δ_Q and heat transfer δ_T. Also in Figure 3.5, it is evident that δ contains errors that have accumulated through the analysis between $t = 0$ and $t = n\Delta t_{AT}$. Isolating error contributions from both δ_Q and δ_T and at each time step n has been identified as a goal of this chapter, and the subsequent step-wise and partitioned model discrepancies are developed in Sections 3.4.1 and 3.4.2 and compared against the global discrepancy treatment in Section 3.5.

3.4.2 Step-wise Model Discrepancy

A step-wise model discrepancy approach is depicted in Figure 3.6 where a model discrepancy δ_n is applied at each time instant n. Again, the coupled model prediction is treated as a black box,

but over each time step. Formulated in Eq. (3.6), the corrected predictions $T_{w,n}^{*}$ at each time step feed forward to both the next prediction and model discrepancy. This approach has been shown to have additional calibration advantages in earlier work [64] where a linear step-wise discrepancy model achieved the same prediction order as a quadratic global discrepancy model. Therefore, fewer parameters were required for calibration in the step-wise approach.

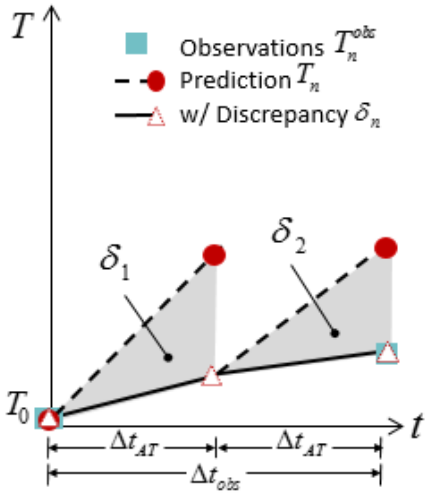

Figure 3.6. Step-wise model discrepancy through time

$$T_{w,n}^{*} = T_{w,n}(T_{w,n-1}^{*}, Q_t) + \delta_n(T_{w,n-1}^{*}, Q_n) \tag{3.6}$$

However, while the step-wise formulation isolates errors that aggregate between time steps, it does not attribute them to either $\delta_{Q,n}$ or $\delta_{T,n}$. Therefore, the step-wise discrepancy formulation treats the temperature prediction at each time step like a black box. In contrast, a partitioned model discrepancy approach developed in Section 3.4.3 fully resolves the error contributions from each model and time-step in the partitioned analysis.

3.4.3 Partitioned Model Discrepancy

The partitioned model discrepancy approach is demonstrated in Figure 3.7 where both heat flux and temperature predictions are corrected at each time instant n with partitioned model discrepancies $\delta_{Q,n}$ and $\delta_{T,n}$. In this way, $\delta_{Q,n}$ at time-step n effectively modifies the temperature

gradient across time step Δt_{AT} while $\delta_{T,n}$ is specifically attributed to the inadequacy of linear heat transfer across that Δt_{AT}. Equations (3.7) and (3.8) show that the corrected wall temperature $T_{w,n}^*$ is a function of $T_{w,n\text{-}1}^*$, as in the step-wise approach, however the corrected heat flux Q_n^* (which is also dependent on $T_{w,n\text{-}1}^*$) is also used for the prediction of $T_{w,n}^*$.

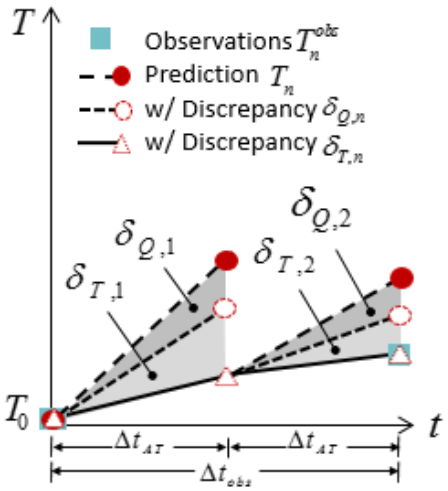

Figure 3.7. Partitioned model discrepancy through time

$$T_{w,n}^* = T_{w,n}(T_{w,n-1}^*, Q_n^*) + \delta_{T,n}(T_{w,n-1}^*, Q_n^*) \tag{3.7}$$

$$Q_n^* = Q_n(T_{w,n-1}^*) + \delta_{Q,n}(T_{w,n-1}^*) \tag{3.8}$$

The three input-dependent model discrepancy resolutions (global, step-wise, and partitioned) will be calibrated and compared in Section 3.5.

3.5 Model Error Calibration in Time-Dependent, Coupled Analyses

Expanding the static Bayesian network from Chapter 2 to the dynamic Bayesian network (DBN) shown in Figure 3.8, the quasi-static aerothermal predictions from Section 3.2 and the temperature data from Section 3.3 are used for calibrating model discrepancy calibration through time. Specifically, Figure 3.8 depicts a DBN for the partitioned model discrepancy approach when two sources of model uncertainty – δ_Q and δ_T are propagated through the network at each time-

step. The data interval Δt_{obs} and aerothermal prediction interval Δt_{AT} need not be equivalent as shown in Figure 3.8 (see Chapter 5), but are chosen as such for the calibrations in this chapter.

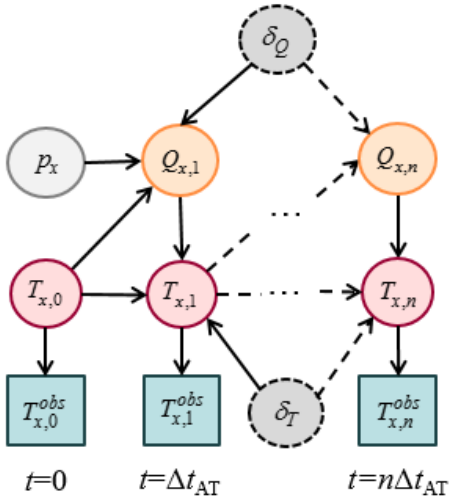

Figure 3.8. Dynamic Bayesian network for partitioned model discrepancy

The corresponding DBN for the step-wise discrepancy approach removes the node for δ_Q and replaces δ_T in Figure 3.8 with a single discrepancy term δ_n applied at each time step n. The global discrepancy approach, however, removes the transient aspects of the DBN entirely and is therefore is equivalent to a static Bayesian network with five nodes: inputs p_x and $T_{x,0,}$ model prediction vector $\boldsymbol{T}_x$ over all time-steps n, data vector $\boldsymbol{T}_x^{obs}$, and model discrepancy δ.

Parameterizing each model discrepancy formulation is discussed in Section 3.5.1 along with the prior distributions of the model discrepancy terms. The posterior temperature predictions for the front, middle, and back points of the Run 30 dome are compared in Section 3.5.2.

3.5.1 Model Discrepancy Parameterization and Prior Distributions

Since input-dependent error models are desired [62], Eqs. (3.5) through (3.8) each show that the model discrepancy term has the same input variables as the model itself. This implies that for coupled analyses, model discrepancies should likewise be coupled. This warrants exploration into appropriate discrepancy parameterization for such time-dependent analyses where the individual

models in a simulation are not functions of time but of the time-step and model outputs from a) the same model at previous time-instances, or b) other models at the either the current or previous time instants.

Figure 3.9 shows the relationship between the nominal temperature error compared to a ratio of heat flux $Q_{x,n}$ and a reference heat flux $Q_{x,ref}$, that is, the initial heat flux in the calibration region $Q_{x,0}$ at t = 1s. This heat flux ratio serves the following two purposes: 1) during calibration the ratio $Q_{x,n}/Q_{x,0}$ is used to model the error as a function of the distance between $Q_{x,n}$ and the initial condition $Q_{x,0}$, and 2) during prediction, the heat flux ratio $Q_{x,n}/Q_{x,ref}$ provides a connection between the calibration region and prediction where an increasing heat flux ratio increases the uncertainty in the model error. The same concept can be applied using a temperature ratio $T_{x,n}/T_{x,0}$ during calibration and $T_{x,n}/T_{x,ref}$ for prediction.

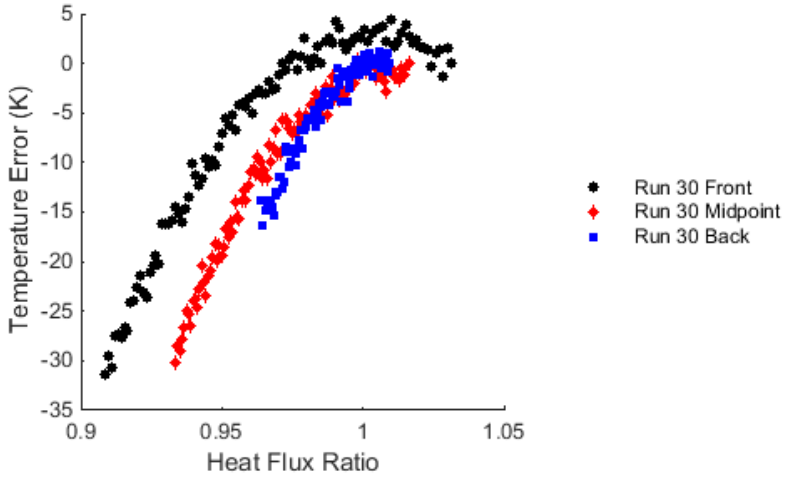

Figure 3.9. Temperature errors vs. heat flux ratio

Therefore, the proposed discrepancy parameterizations for the global, step-wise, and partitioned discrepancy approaches are shown in Eqs. (3.9), Eq. (3.10), and Eqs. (3.11) through (3.12), respectively. Since Figure 3.9 displays a quadratic trend with respect to heat flux ratio, the global

model discrepancy in Eq. (3.9) is chosen to be quadratic with respect to $Q_{x,n}/Q_{x,ref}$ and linear in $T_{x,n}/T_{x,ref}$.

$$\delta = a_0 + a_1\left(\frac{Q_{x,n}}{Q_{x,ref}}\right) + a_2\left(\frac{Q_{x,n}}{Q_{x,ref}}\right)^2 + a_3\left(\frac{T_{x,n}}{T_{x,ref}}\right) \tag{3.9}$$

As previously discussed, it was shown in [64] that a step-wise approach could use a lower order discrepancy model to obtain the same prediction order as the global approach. Therefore, the step-wise discrepancy parameterization in Eq. (3.10) has a linear relationship to heat flux ratio rather than quadratic shown in Eq. (3.9).

$$\delta_n = b_0 + b_1\left(\frac{Q_{x,n}}{Q_{x,ref}}\right) + b_2\left(\frac{T_{x,n}}{T_{x,ref}}\right) \tag{3.10}$$

Finally, the proposed partitioned model discrepancy formulations arise by further subdividing the step-wise discrepancy formulation Eq. (3.10) into $\delta_{T,n}(Q_n^*)$ and $\delta_{Q,n}(T_{n-1}^*)$ contributions. Note that $\delta_{T,n}(Q_n^*)$ differs from $\delta_{T,n}(T_{n-1}^*, Q_n^*)$ in (3.7), however, the error function in Eq. (3.11) is deemed appropriate for partitioned approaches since the goal of calibrating model discrepancy at the partitioned resolution is to remove the error contributions from previous time instants or models in the analysis.

$$\delta_{T,n} = d_0 + d_1\left(\frac{Q_{x,n}}{Q_{x,ref}}\right) \tag{3.11}$$

$$\delta_{Q,n} = c_0 + c_1\left(\frac{T_{x,n}}{T_{x,ref}}\right) \tag{3.12}$$

Prior distributions for the discrepancy coefficients for each model discrepancy resolution are considered uniform random variables with lower and upper bounds shown in Table 3.2. The prior

distribution bounds for each calibration parameter were chosen by passing midpoint values of the other parameters within the same discrepancy model through the partitioned analysis and comparing the nominal temperature predictions to the Glass and Hunt data.

Table 3.2. Uniform prior distributions for discrepancy model parameters

Discrepancy Resolution	Parameter	Lower Bound	Upper Bound
Global (Eq. (3.9))	a_0 (K)	-700	-400
	a_1 (K)	2e3	5e3
	a_2 (K)	-3e3	-1.5e3
	a_3 (K)	-900	-400
Step-Wise (Eq. (3.10))	b_1 (K)	-2	5
	b_2 (K)	-4	15
	b_3 (K)	-20	5
Partitioned (Eq. (3.11) and Eq. (3.12))	c_0 (W/cm^2)	-5e5	5e5
	c_1 (W/cm^2)	-5e5	5e5
	d_0 (K)	-20	20
	d_1 (K)	-50	50

The uncertain model discrepancy coefficients shown in Table 3.2 are calibrated in Section 3.5.2 using time-dependent temperature data from Runs 30, 31, 32.

3.5.2 Calibration and Prediction Confidence Assessment

Recall that the temperature histories were subdivided into calibration and validation sets where the temperatures observed from 1 to 4 seconds are used for calibration. In Section 3.3, the nonlinearity within this region provided the initial justification for calibrating with this data set, however, the additional benefits to calibration are discussed next.

Because the calibration set of temperatures starts at t = 1s, the reference wall temperature $T_{x,ref}$ is non-uniform which was initially observed to capture the spatial differences between model errors through time and increased uncertainty reduction in the temperature ratio coefficients (i.e., a_4, b_2, and d_1). A non-uniform heat flux distribution for calibration was not of concern, however, because $Q_{x,ref}$ varies across the dome inherently from its relationship with aerodynamic pressure

and dome slope. In addition, the effects of extrapolation can be investigated when $T_{x,n}$ is less than $T_{x,ref}$ (between t = 0 and 1s) and greater than $T_{x,ref}$ (when t = 4 to 5s).

Bayesian model calibration (see Section 2.2) was performed over the DBN using 10e3 slice samples for each discrepancy resolution across the 3 second calibration temperature history (t = 2 to 4s) from Runs 30, 31, and 32 domes.

The post-calibration wall temperature predictions are compared against full 5 second temperature histories (initial wall temperature $T_{x,0}$ = 300K) at the Run 30 dome front, midpoint, and back follow from each discrepancy approach in Figure 3.10, Figure 3.11, and Figure 3.12. This means that the initial conditions and analysis durations differed between calibration and validation. Figure 3.10, Figure 3.11, and Figure 3.12 compare the predictions from calibrated aerothermal models from the global, step-wise, and partitioned discrepancy approaches, respectively.

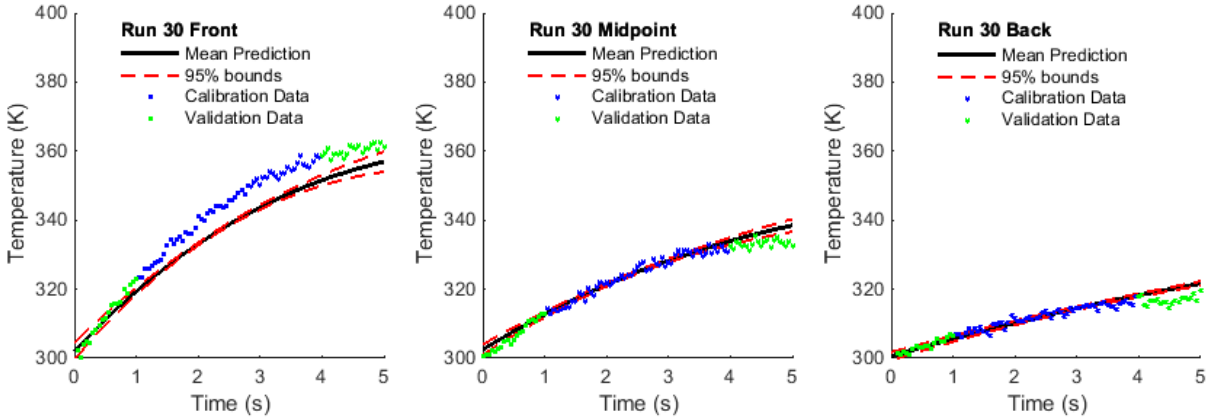

Figure 3.10. Post-calibration wall temperature predictions with a global model discrepancy resolution

At the front of the Run 30, higher heat fluxes, temperatures, and nonlinearities are observed in the data. Posterior predictions with the global model discrepancy approach in Figure 3.10 has large prediction uncertainty in the validation regions. Furthermore, the prediction from the global approach has significant prediction bias after t = 1s. At the dome midpoint and back locations in

Figure 3.10, the there is less uncertainty about the prediction since the heat flux and temperature ratios are small in comparison to those at the dome front and thus the propagating less uncertainty from the calibrated discrepancy coefficients.

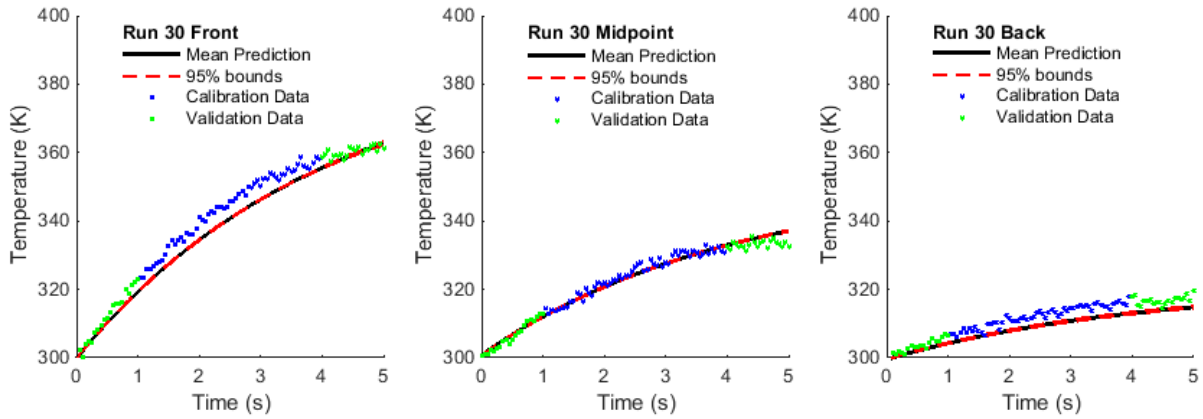

Figure 3.11. Post-calibration temperature predictions with a step-wise model discrepancy resolution

In Figure 3.11, the step-wise model discrepancy approach demonstrates less prediction uncertainty and bias at the front of the Run 30 dome than the global approach. The step-wise approach also predicts less uncertainty at the midpoint and the back of the dome, where it also demonstrates a prediction order increase over the global approach by more closely following the trend of the temperature history at the back of the Run 30 dome.

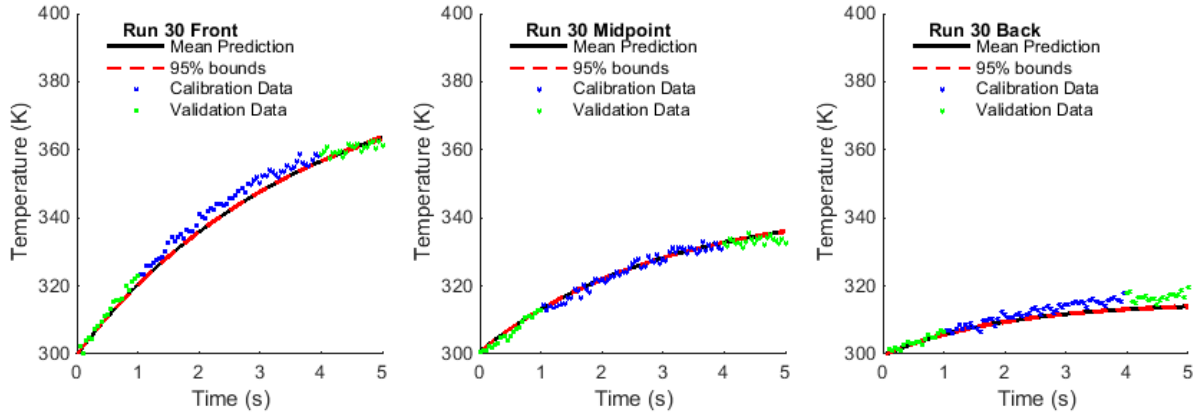

Figure 3.12. Post-calibration wall temperature predictions with a partitioned discrepancy resolution

Finally, Figure 3.12 shows the post-calibration predictions using a partitioned model discrepancy resolution for the aerothermal prediction. By inspection, the partitioned approach has reduced the prediction bias at the front of the Run 30 dome in the nonlinear region (1 to 4s) compared to the step-wise approach in Figure 3.11.

Because both prediction bias and uncertainty are of interest, model reliability metric [65] is employed for quantitative comparisons between the posterior predictions. Equation (3.13) shows the reliability $R(t)$ at time t is computed as the probability that the difference between the stochastic model prediction y and validation data $y_D(t)$ is within a pre-specifed tolerance τ. [66]. When the tolerance is chosen as the measurement error σ_D (assuming Gaussian), the model is equivalent to computing the area of overlap between the prediction and measurement distributions. A temperature tolerance of 5 K is used to compute the prediction reliabilities from the global, step-wise, and partitioned discrepancy approaches that are presented in Figure 3.13.

$$R(t) = \mathrm{P}[\tau < y(t) - y_D(t) < \tau] \tag{3.13}$$

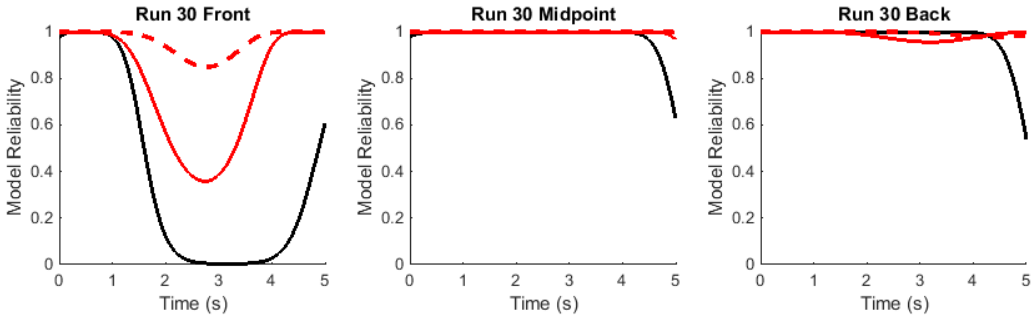

Figure 3.13. Posterior prediction reliability across Run 30 dome using global (black), step-wise (solid red), and partitioned (dashed red) discrepancy approaches

Figure 3.13 demonstrates that, for a prediction reliability tolerance of 5 K, the partitioned discrepancy approach maintains a high level of reliability across the front of the Run 30 dome. In the nonlinear region between (t = 1 to 4s) the prediction reliability decreases to 0.8 and quickly recovers to 1. The prediction reliability from the step-wise approach also recovers to 1 after the

nonlinear region, however, the added bias affects the model reliability more severely. Finally, the global model discrepancy approach also recovers prediction reliability after t = 4s due to increasing uncertainty the extrapolation in that region that is superficially inflating the probability of agreement with the data and loses extrapolation ability quickly after t = 4s at the midpoint and back locations.

3.6 Conclusion

Geared toward quantifying model error aggregation in coupled, time-dependent analyses, this chapter developed and compared three model discrepancy resolutions (i.e., global, step-wise, and partitioned) for integration into the Bayesian model calibration framework. It was shown that the partitioned error formulation isolated the model error contributions om coupled aerodynamic heating and 1D heat transfer predictions through time using one source of time-dependent temperature data. In addition, the posterior predictions with the model discrepancies from the partitioned approach showed increased confidence in both extrapolation and nonlinear regions compared to the step-wise and global approaches.

Effective parameterization of the model error for each discrepancy resolution was addressed, where the step-wise and partitioned model discrepancies were selected as lower order than those from the global approach. While this did not result in fewer calibration parameters between the global and partitioned approaches for the application (each had 4 parameters), it did result in more intuitive selection of prior distributions before calibration of the partitioned discrepancy parameters as well as interpretation of the posterior results. Closely tied to parameterizing the model discrepancy was the choice of calibration data, where choosing a non-uniform initial temperature distribution better captured the spatial differences among error aggregation rates through time.

One highlight from the comparison between the three model discrepancy resolutions is the lack of additional computational expense in the application the step-wise and partitioned model discrepancies through the dynamic Bayesian network compared to the global resolution. The only differences in computational effort between the three methods were seen during calibration due to the number of calibration parameters. In effect, the calibration with step-wise model discrepancy would have converged faster with three calibration parameters than the global or partitioned approaches with four, however, the 40% gain in prediction confidence through the nonlinear region using the partitioned model discrepancy approach far outweighs the added calibration expense.

Further investigation into the impact of propagating uncertainty from the stochastic aerodynamic pressure predictions through the dynamic aerothermal Bayesian network is needed. Other guidance and procedures for simplifying recalibration over dynamic Bayesian networks when new data is available for integration is also of interest, for example, if data becomes available on heat flux through time to use in an aerothermal Bayesian network.

CHAPTER 4

EFFICIENT GLOBAL SENSITIVITY ANALYSIS FOR TRANSIENT, MULTIDISCIPLINARY PROBLEMS

4.1 Introduction

Sensitivity analyses provide useful insights into both the forward problem of uncertainty propagation as well as the inverse problem of uncertainty reduction by quantifying how the prediction is influenced by individual uncertainty sources. Until recently, sensitivity analyses conducted for forward problems of uncertainty quantification (UQ) in multidisciplinary analyses (MDA) and multidisciplinary design optimization (MDO) have focused on gradient-based sensitivities which are local. The inclusion of uncertainty in MDA has led to significant research in robust MDO approaches require quantifying how the system performance is affected by individual sources of uncertainty [12–15] using global sensitivity analysis. In contrast to gradient-based local sensitivities at a chosen nominal value, global sensitivity analyses (GSA) use the entire probability distribution of the uncertainty sources.

Sobol' indices are a set of variance-based global sensitivity indices that quantify both individual and interactive effects between uncertain parameters on the model output variance [16]. The interactive effects are especially important in multidisciplinary simulations where parameter interactions among multiple disciplinary models may contribute to additional uncertainty in the predicted response [17]. Sobol' indices for GSA present computational challenges, however, due to the large number of input-output samples needed to estimate the uncertainty contributions from even a single variable and quickly become intractable as the number of uncertain parameters increases. Saltelli's widely-used class of matrix methods [18] have improved the computational

efficiency, however, the computational burden may still be too large in problems of high-dimension and in the presence of expensive models.

To alleviate the computational burden of GSA, surrogate models have been used in place of the expensive simulations [19,20]. Emphasis on non-intrusive surrogate modeling methods for uncertainty propagation have been seen for aero-applications. Lamorte et al. investigated the implementation of a stochastic collocation approach for differential equations using polynomial response surfaces to propagate uncertainty due to onset of transition and unstable boundary layers [21–23]. Hosder used a stochastic response surface obtained with non-intrusive polynomial chaos models for both uncertainty propagation and simplified sensitivity calculations [24]. However, the additional step of training and validating a surrogate to get accurate predictions is not straight-forward and is dependent on the set of training points and the dimension of the input space. Hu and Mahadevan [25] recently investigated global sensitivity analysis-enhanced surrogate (GSAS) modeling for reliability analysis, where the new training points are selected based on GSA results, however, this further motivates more efficient GSA computations that are not reliant on a surrogate model.

Other work has focused on efficiently using existing sets of input-output samples from model verification, validation, or calibration stages of model development. For example, Li and Mahadevan developed a modular global sensitivity analysis (MGSA) methodology that uses stratified sampling to assign uniform weights to one-dimensional strata to compute first order Sobol' indices. A similar idea has been explored using an importance sampling-based kernel regression method (ISK-GSA) developed by Sparkman et. al. [26] where the choice of kernel need not be uniform and can follow the distribution type of the parameter. Furthermore, the ISK-GSA method calculates Sobol' indices in fewer model runs and has the desired capability of additionally

being able to efficiently compute the total effects Sobol' indices using the same existing input-output samples. Thus, the ISK-GSA methodology provides the foundation for the research contributions presented in this chapter.

For inverse problems in UQ, GSA is used to identify the subset of uncertain inputs and parameters that are candidates for uncertainty reduction through calibration. Calibration requires a significant number of model evaluations within the likelihood until the posterior distribution convergence, however, these input-output samples cannot be integrated into existing methodologies since the posterior distributions of the parameters exhibit correlation. In the literature, GSA with dependent variables are typically consigned to either expensive double-loop computations, grouping the correlated parameters into one auxiliary variable [34], or a combination of both approaches. As such, the input-output samples from calibration have yet to be used efficiently in post-calibration GSA and is a significant research gap that is addressed with the methodology developed herein.

First, the proposed methodology addresses independent variables and improves the ISK-GSA methodology developed by Sparkman et. al. in Ref. [26] by using quasi-random sequences. Studies have found that quasi-random sequences provide optimal space-filling designs in higher dimensions [35] and the combined methodologies improve the convergence in estimating the sensitivity of a model output to independent sources of uncertainty. Second, the ISK-GSA method is generalized to allow consideration of dependent variables which are observed a) in posterior samples of calibrated parameters and b) among parameters and model outputs passing between coupling models in MDA.

The outline of this chapter is as follows: Sobol' indices for global sensitivity analyses are presented in Section 4.2 followed by an overview of the ISK-GSA method as developed by

Sparkman et al. in 4.3. Section 4.4 presents the proposed methodology suitable for independent variables that uses space-filling quasi-random number generators together with the ISK-GSA method. In Section 4.5, the generalized ISK-GSA methodology is developed to handle dependent variables which are observed after calibration and between coupled models through time and is demonstrated on an illustrative time-dependent example in Section 4.6. Section 4.7 applies the generalized ISK-GSA methodology to compute the posterior GSA results through time for a coupled, time-dependent aerothermal analysis and demonstrates the effects of model coupling and parameter correlation on the sensitivity estimates.

4.2 Sobol' Sensitivity indices

After propagating D input sources of uncertainty $\boldsymbol{X}_{1xD}$ to a model output Y, the variance decomposition theorem in Eq. (3.14) states that the total variance of the model output $Var(Y)$ can be decomposed into the summation of a) the variance of the expectation of model output conditioned on the d^{th} input variable X^d (d = 1 to D) with all other variables $\boldsymbol{X}^{\sim d}$ varying and b) the expectation of the variance of Y conditioned the same the same set [70,72]. The summation of these two components are shown in Eq. (3.14).

$$Var(Y) = Var_{X^d}\{E_{\boldsymbol{X}^{\sim d}}[Y \mid X^d]\} + E_{X^d}\{Var_{\boldsymbol{X}^{\sim d}}[Y \mid X^d]\} \tag{3.14}$$

The first-order Sobol' index $S_{d,1}$ shown in Eq. (3.15) is the ratio of variance contributed by X^d to the total variance in Eq. (3.14). The most straightforward and widely-used procedure to compute Sobol' indices is the double-loop (DL) method. In the double-loop method, computing the first-order Sobol' index of variable X^d in Eq. (3.15) first consists of an outer-loop each at a fixed input x_i^d (i = 1 to N_{outer}), and an inner-loop over $\boldsymbol{x}_j^{\sim d}$ (j = 1 to N_{inner}). Within the inner loop, an estimate of the conditional expectation $E_{\boldsymbol{X}^{\sim d}}[Y \mid X^d = x_i^d]$ is obtained using N_{inner} samples of $\boldsymbol{X}^{\sim d}$. This process

is repeated by resampling X^d for each of N_{outer} outer-loop iteration and the variance of the conditional expectations from each choice of x_i^d is used in Eq. (3.15).

$$S_{d,1} = \frac{Var_{X^d}\{E_{\boldsymbol{X}^{\sim d}}[Y \mid X^d]\}}{Var(Y)} \tag{3.15}$$

The sum of the first-order indices across all D dimensions is equal to 1 (within a margin of finite sampling error), or less than 1 if parameter interactions are significant which are not captured in the first-order Sobol' index. Therefore, the total effects Sobol' index $S_{d,T}$ inherently accounts for the first-order effects of X^d as well as the effects of interactions between X^d and all other variables $\boldsymbol{X}^{\sim d}$. Parameter interactions may either be a result of nonlinearities intrinsic to the model, statistical correlations between the parameters in the input space, or a combination of both. The total effects indices in Eq. (3.16) are derived by dividing the second term in Eq. (3.14) by $Var(Y)$.

$$S_{d,T} = \frac{E_{\boldsymbol{X}^{\sim d}}\{Var_{\boldsymbol{X}^d}[Y \mid X^{\sim d}]\}}{Var(Y)} = 1 - \frac{Var_{\boldsymbol{X}^{\sim d}}\{E_{X^d}[Y \mid \boldsymbol{X}^{\sim d}]\}}{Var(Y)} \tag{3.16}$$

Using the double-loop method and $N_{outer} = N_{inner} = N$, the number of model evaluations M required for estimating both first-order and total effect Sobol' indices for D parameters is on the order of $O(2DN^2)$), which quickly become intractable. These computational expenses were addressed by Saltelli et al. [72] where a class of matrix-column exchange (MCE) methods were developed to simplify GSA by approximating the variances in the numerators In contrast to the double-loop method for computing the indices, MCE methods require $M \sim O(2(2+D)N)$ because they estimate the variances in the numerators of. First, an initial input matrix of dimension $2N$xD is sub-divided into two NxD matrices $\boldsymbol{X}_A$ and $\boldsymbol{X}_B$. These two input matrices are then propagated through the model for y_A and y_B. Next, a matrix $\boldsymbol{X}_{AB}^d$ for each dimension d is then formed by permuting the d^{th} column of $\boldsymbol{X}_A$ with the d^{th} column of $\boldsymbol{X}_B$ such that $\boldsymbol{X}_{AB}^d = [\boldsymbol{X}_A^1, \ldots \boldsymbol{X}_A^{d-1}, \boldsymbol{X}_B^d, \ldots,$

$\boldsymbol{X}_A^D$]. Each $\boldsymbol{X}_A$, $\boldsymbol{X}_B$, and $\boldsymbol{X}_{AB}^d$ (d = 1 to D**)** are propagated through the model for output matrices y_A, y_B, and y_{AB}^d (d = 1 to D), respectively. Several variance formulations have been explored to use MCE methods for Sobol' indices, and the approaches from [80] are presented below in Eqs. (3.17) and (3.18).

$$Var_{X^d}\{E_{\boldsymbol{X}^{\sim d}}[Y \mid X^d]\} = \frac{1}{N-1}\sum_{n=1}^{N} y_{A,n} y_{BA,n}^d - \frac{1}{N}\sum_{n=1}^{N} y_{A,n} y_{B,n} \tag{3.17}$$

$$Var_{\boldsymbol{X}^{\sim d}}\{E_{X^d}[Y \mid \boldsymbol{X}^{\sim d}]\} = \frac{1}{N-1}\sum_{n=1}^{N} y_{A,n} y_{AB,n}^d - \frac{1}{N}\sum_{n=1}^{N} (y_{A,n})^2 \tag{3.18}$$

The challenge with computing Sobol' indices for time-dependent, multidisciplinary analyses with either the DL or MCE methods is two-fold; 1) convergence of the sensitivity estimate is slow from sub-optimal sampling and expensive models, and 2) MCE methods, among others, are unable to accommodate correlated quantities. First, computational effort is addressed using ISK-GSA in its original form and improved upon using Sobol' sequences in Section 4.3. Performing post-calibration GSA with correlated parameters is addressed in Section 4.3.

4.3 Importance Sampling-based Kernel Regression Estimator for Sobol' Indices (ISK-GSA)

First, the independent concepts of kernel regression and importance sampling are introduced in Sections 4.3.1 and 4.3.2 and then used together with the ISK-GSA methodology in Section 4.3.3 to estimate Sobol' indices. The ISK-GSA methodology is then compared against DL and MCE methods using a simple example presented in Section 4.3.4 to demonstrate the computational savings offered by ISK-GSA and serve as a benchmark for the specific improvements to GSA with independent variables shown in Section 4.4. It also serves as the foundation for the expansion to GSA with correlated variables in Section 4.5.

Consider M available model evaluations from various stages of model development with input-output relationships X_{MxD} and Y_{Mx1}. For each model evaluation m = 1 to M and input dimension d = 1 to D, an estimate of the conditional expectation $E_{X\sim d}[Y|X^d = x_m{}^d]$ is achieved by weighting each model output Y as shown in Eq. (3.19).

$$E_{X^{\sim d}}[Y \mid X^d = x_m^d] = \sum_{j=1}^{M} y_j w_j^d \tag{3.19}$$

Kernel regression, importance sampling, and ISK-GSA each use Eq. (3.19) to estimate the conditional expectation of Y given $x_m{}^d$, yet differ in how each of the M available model outputs are weighted.

4.3.1 Kernel regression weights

Kernel regression uses a locally weighted average within the neighborhood of $x_m{}^d$ where neighborhood is defined using a kernel function $K_d(x_m{}^d\text{-}x_j{}^d)$ centered around on $x_m{}^d$. The kernel function outputs are normalized according to Eq. (3.20) to define the weights $w_j{}^d$ for kernel regression in Eq. (3.19).

$$w_j^d = \frac{K_d(x_m^d - x_j^d)}{\sum_{j=1}^{M} K_d(x_m^d - x_j^d)} \tag{3.20}$$

Kernel functions can be a uniform or constant for all $x_j{}^d$ within a certain Euclidean distance h_d from $x_m{}^d$ or parametrized as continuous functions that decrease over the distance between the kernel center $x_m{}^d$ and $x_j{}^d$. In this work, Gaussian kernels in each dimension are defined with kernel center $c_d = x_m{}^d$ and bandwidth $h_d = 1.06\sigma_d M^{-1/5}$ as suggested for Gaussian kernels in [21] where σ_d is standard deviation of the parameter X^d. Thus, the kernel bandwidth in each dimension d scales with input variance σ_d and decreases as the number of model evaluations M increases implying

that the conditional expectation estimate in Eq. (3.19) becomes more localized. These local conditional expectation estimates within the global probability space are supplemented with concepts from importance sampling in the ISK-GSA methodology and are discussed next.

4.3.2 Importance sampling weights

Importance sampling is a variance reduction technique that first distinguishes between the sampling density $S_D(\boldsymbol{x})$ of the input space from which M model evaluations are generated and a target density $T_D(\boldsymbol{x})$ from which statistics of the output are desired. In engineering reliability analyses, importance sampling defines the sampling density as a region of the input space with a high probability of an event of interest (e.g., failure) [40]. Therefore, the target probability of the event over the entire set of possible inputs reweights the outcomes observed in the sampling region with the probability of that combination of inputs.

For the kernel regression weights, however, the sampling density $S_d(x^d)$ and its statistics are known and the local statistics with a smaller subspace of the inputs are to be estimated. Therefore, for each input dimension d, an estimate the output within a target region $T_d(x^d)$ using samples from the sampling density $S_d(x^d)$ is achieved using normalized importance-sampling weights w_j^d defined according to Eq. (3.21).

$$w_j^d = \frac{T_d(x_j^d) / S_d(x_j^d)}{\sum_{j=1}^{M} T_d(x_j^d) / S_d(x_j^d)} \tag{3.21}$$

4.3.3 Importance sampling-based kernel regression weights

The formulation for the importance sampling-based weights for ISK-GSA arises out of equating the kernel function $K_d(x_m^d - x_j^d)$ in (3.20) that defines the neighborhood around x_m^d and the marginal target density $T_d(\boldsymbol{x})$ from Eq. (3.21). Likewise, and the sampling density $S_d(x^d)$ is equivalent to marginal probability density of the available inputs $p_d(x^d)$ in each input dimension. Thus, for each

model evaluation *M*, the conditional expectation in Eq. (3.19) for first order effects is estimated using the ISK weights in Eq. (3.22).

$$w_j^d = \frac{K_d(x_m^d - x_j^d) / p_d(x_j^d)}{\sum_{j=1}^{M} K_d(x_m^d - x_j^d) / p_d(x_j^d)} \tag{3.22}$$

The conditional expectation for total effects index in Eq. (3.23) is computed using Eq. (3.24) for the normalized weights. The normalized weight in Eq. (3.24) defines the neighborhood around $\boldsymbol{x}_m^{\sim d}$ with a *D*-1 dimensional kernel $K_{\sim d}(\boldsymbol{x}_m^{\sim d} - \boldsymbol{x}_j^{\sim d})$. The probability density $p_{\sim d}(\boldsymbol{x}_m^{\sim d})$ assumes that the *d* and ~*d* distributions are are independent.

$$E_{X^d}[Y \mid X^{\sim d} = x_m^{\sim d}] = \sum_{j=1}^{M} y_j w_j^{\sim d} \tag{3.23}$$

$$w_j^{\sim d} = \frac{K_{\sim d}(x_m^{\sim d} - x_j^{\sim d}) / p_{\sim d}(x_j^{\sim d})}{\sum_{j=1}^{M} K_{\sim d}(x_m^{\sim d} - x_j^d) / p_{\sim d}(x_j^{\sim d})} \tag{3.24}$$

The following section uses the ISK-GSA methodology to compute the sensitivities for a simple mathematical example.

4.3.4 Mathematical Example using ISK-GSA

To demonstrate the computational benefits of the ISK-GSA methodology for sensitivity analysis, consider a two-parameter model $y = 2x_1 + x_2$ where both parameters x_1 and x_2 are considered to be independent Gaussian random variables with zero mean and unit variance. The analytical first order and total effect Sobol' indices are presented in Eqs. (3.25) and (3.26) and are shown to be equivalent since there are neither parameter interactions within the model nor correlations.

$$S_{x_1,1} = S_{x_1,T} = \frac{Var_{x_1}\{E_{x_2}[y \mid x_1]\}}{Var(y)} = \frac{4\sigma_{x1}^2}{4\sigma_{x1}^2 + \sigma_{x2}^2} \tag{3.25}$$

$$S_{x_2,1} = S_{x_2,T} = \frac{Var_{x_2}\{E_{x_1}[y \mid \theta_2]\}}{Var(y)} = \frac{\sigma_2^2}{c^2\sigma_1^2 + \sigma_2^2} \tag{3.26}$$

Using Eqs. (3.25) and (3.26) and $\sigma_{x1} = \sigma_{x2} = 1$ the true first-order and total effects indices of x_1 and x_2 are 0.8 and 0.2, respectively. Figure 4.1 compares sensitivity estimates from N iterations of sensitivities from three methods: the double-loop (DL) method, Saltelli's matrix column-exchange method (MCE), and the ISK-GSA method. Latin-hypercube sampling (LHS) was used to randomly sample the $D = 2$ dimensional input space for $N = 1$ to 400 sensitivity iterations. The number of model evaluations required for N sensitivity iterations of each of the three methods compared is shown in the legend of Figure 4.1.

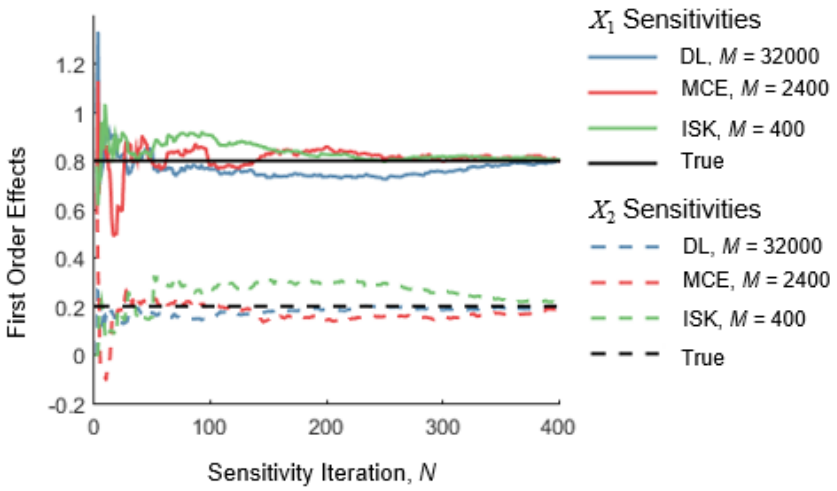

Figure 4.1. First order effects of x_1 (top) and x_2 (bottom) across N sensitivity iterations

The legend in Figure 4.1 indicates that the number of model evaluations needed to compute the first order and total effect indices for the ISK-GSA method at each sensitivity iteration N is $O(N)$ compared to $O(2DN^2)$) and $O(2(2+D)N)$ for the DL and MCE methods, respectively. In effect,

these results show that the ISK-GSA method requires less than 1% of the model evaluations required for the DL estimate and less than 17% compared to the MCE method.

Furthermore, the ISK-GSA methodology can update the sensitivity estimate after each model evaluation, meaning the convergence of the sensitivity estimates may be monitored with each successive evaluation of the model. This example will be used in Section 4.4 to demonstrate improved convergence properties when the ISK-GSA methodology employs quasi-random number generators as opposed to LHS. These improvements are especially useful during pre-calibration sensitivity analyses where the analyst is in control of the sampling space. Then, extension of the ISK-GSA methodology to correlated quantities is derived for post-calibration sensitivity analysis in Section 4.5 and applied to the posterior sensitivities the coupled aerothermal models through time in Section 4.6.

4.4 Sensitivity Analysis for Independent Variables with ISK-GSA and Quasi- Random Number Generators

Quasi-random and pseudo-random number generators are two distinct classes of algorithms used to generate random Monte Carlo samples for numerical integration, uncertainty propagation, and sensitivity analysis. Pseudo-random number generators (e.g., latin-hypercube) have the advantages of being most random-like, as shown in the left-hand side of Figure 4.2, but result in regions of high and low density that can prolong convergence of the integration. Quasi-random sequences (e.g., Sobol' sequences [81], Halton sequences [82], or Hammersley sets [83]) were developed for even coverage over a high-dimensional integration domains and are guaranteed to cover the domain of interest evenly which results in improved efficiency in numerical integration schemes [81].

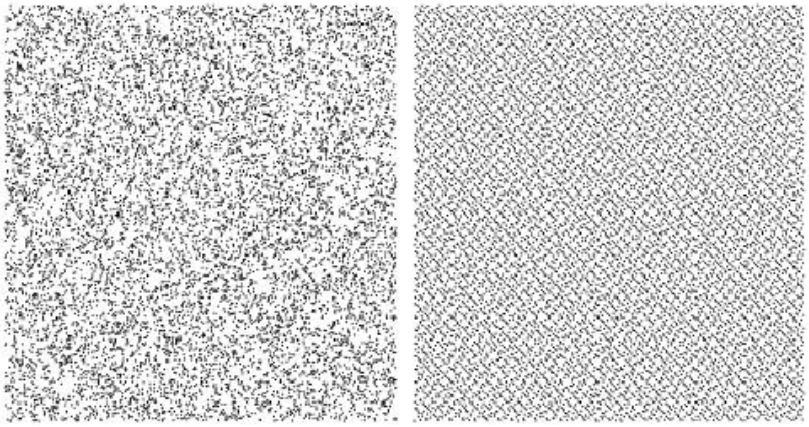

Figure 4.2. Latin-Hypercube design (left) and Sobol' sequence (right) with 10,000 points

The proposed methodology uses the Sobol' sequence to fill the input space sequentially after each iteration $n = 1$ to N where the difference in the domain coverage between $(n\text{-}1)^{th}$ and n^{th} iterations is low. Again considering the model $y = 2x_1 + x_2$ from Section 4.3.4, the advantage of using a Sobol' sequence is evident when N is constrained to a low number of model evaluations ($N = 450$) as shown in Figure 4.3. The asymptotic convergence of the sensitivity estimate is observed when a quasi-random Sobol' sequence is used with the ISK-GSA method compared to using a Sobol' sequence with the DL or MCE methods. Thus, a convergence criterion may be applied to terminate the sensitivity analysis and minimize the computational effort. Also, the Sobol' algorithm to sequentially and uniformly fill the space contributes to the 'stair-step' pattern of convergence in Figure 4.3. In particular, the convergence property is observed using the ISK-GSA method as the number of model evaluations M increases and the kernel bandwidth h_d in each dimension $d = 1,\ldots,D$ decreases according to to $h_d = 1.06\sigma_d M^{-1/5}$.

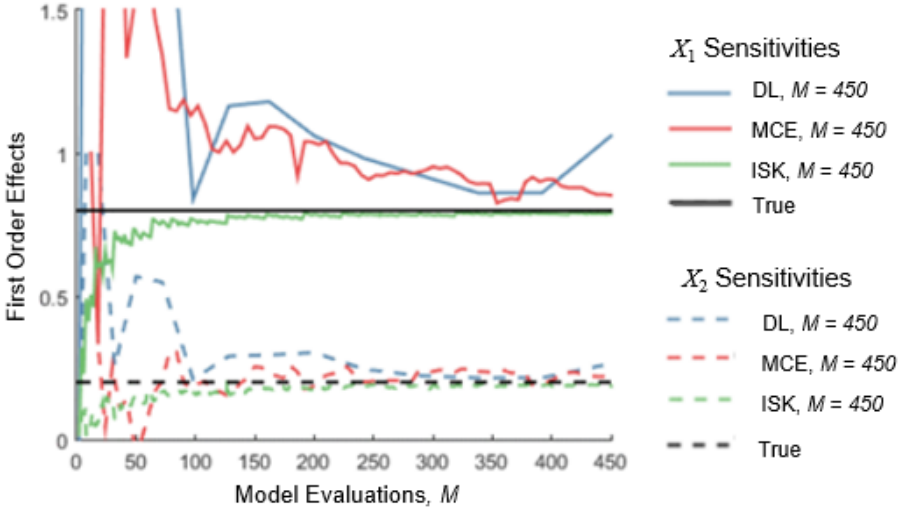

Figure 4.3 Comparison of first order sensitivities using a 2-D Sobol' sequence and LHS design

Global sensitivity analysis for time-dependent, multidisciplinary models will be conducted using the ISK-GSA methodology with Sobol' sequences to reduce the number of coupled simulations required when variables considered independent (e.g., before calibration). This is distinct from post-calibration parameter distributions that are correlated which are addressed by the generalized ISK-GSA methodology in Section 4.5.

4.5 Sensitivity Analysis with Correlated Variables using Generalized ISK-GSA

The derivation of the normalized weights for a generalized ISK-GSA estimate of first-order indices with correlated parameters begins in Eq. (3.27) by recognizing that the first-order ISK-GSA weight in Eq. (3.22) and Eq. (3.27) are equivalent when $\boldsymbol{X}^{\sim d}$ and X^d are independent. In the independent case, the marginal density $p_d(x^d)$ is equivalent to the joint sampling density $p_{\boldsymbol{D}}(\boldsymbol{x})$ divided by the density $p_{\sim d}(\boldsymbol{x}^{\sim d})$. Similarly, a marginal kernel $K_d(x_m^d - x_j^d)$ centered on $\boldsymbol{x}_m^d$ is equivalent to a joint kernel $K_D(\boldsymbol{x}_m^D - \boldsymbol{x}_j^D)$ centered on $\boldsymbol{x}_m^D$ divided by the kernel $K_{\sim d}(\boldsymbol{x}_m^{\sim d} - \boldsymbol{x}_j^{\sim d})$ that is centered on $\boldsymbol{x}_m^{\sim d}$.

$$w_j^d = \frac{K_D(\boldsymbol{x}_m^D - \boldsymbol{x}_j^D)\, p_{\sim d}(\boldsymbol{x}_j^{\sim d}) / K_{\sim d}(\boldsymbol{x}_m^{\sim d} - \boldsymbol{x}_j^{\sim d})\, p_D(\boldsymbol{x}_j^D)}{\sum_{j=1}^{M} K_D(\boldsymbol{x}_m^D - \boldsymbol{x}_m^D)\, p_{\sim d}(\boldsymbol{x}_j^{\sim d}) / K_{\sim d}(\boldsymbol{x}_m^{\sim d} - \boldsymbol{x}_j^{\sim d})\, p_D(\boldsymbol{x}_j^D)} \tag{3.27}$$

Building off Eq. (3.27), the generalized ISK-GSA weight is extended to Eq. (3.28) to apply when X^d and $\boldsymbol{X}^{\sim d}$ are dependent variables . In Eq. (3.28), both the kernel defining the neighborhood around $x_m{}^d$ and probability of $x_j{}^d$ are conditioned on $x_m{}^{\sim d}$. If there is no dependence between X^d and $\boldsymbol{X}^{\sim d}$ then $p_{d|\sim d}(x^d|\boldsymbol{x}^{\sim d}) = p_d(x^d)$ and $K_{d|\sim d}(x^d|\boldsymbol{x}^{\sim d}) = K_d(x^d)$ and Eq. (3.28) reduces to Eq. (3.22).

$$w_j^d = \frac{K_{d|\sim d}(x_m^d - x_j^d \mid X^{\sim d} = x_m^{\sim d}) / p_{d|\sim d}(x_j^d \mid X^{\sim d} = x_m^{\sim d})}{\sum_{j=1}^{M} K_{d|\sim d}(x_m^d - x_j^d \mid X^{\sim d} = x_m^{\sim d}) / p_{d|\sim d}(x_j^d \mid X^{\sim d} = x_m^{\sim d})} \tag{3.28}$$

For completeness, the generalized ISK-GSA weights for total effects indices are presented in Eq. (3.29) and are equivalent to Eq. (3.24) when X^d and $\boldsymbol{X}^{\sim d}$ are independent.

$$w_j^{\sim d} = \frac{K_{\sim d|d}(x_m^{\sim d} - x_j^{\sim d} \mid X^d = x_m^d) / p_{\sim d|d}(x_j^{\sim d} \mid X^d = x_m^d)}{\sum_{j=1}^{M} K_{\sim d|d}(x_m^{\sim d} - x_j^{\sim d} \mid X^d = x_m^d) / p_{\sim d|d}(x_j^{\sim d} \mid X^d = x_m^d)} \tag{3.29}$$

These generalized ISK-GSA methodology is demonstrated on a time-dependent, coupled problem first with independent parameters and then with correlated parameters in Section 4.6. The methodology is then applied to post-calibration sensitivity analysis of coupled aerothermal models in Section 4.7.

4.6 Time-Dependent, Multidisciplinary Example

Consider the models in Eq. (3.30) that are representative of multidisciplinary, time dependent analyses. The models are coupled such 1) the value of model output $y_{1,i,}$ moves from 1 toward zero through time as 2) y_2 increases from the initial condition $y_{2,0}$ at a rate of $y_{1,i}$ across $\Delta t = 0.1s$ until the final time of interest $t_f = 50s$.

$$y_{1,i} = \frac{y_{2,0}}{y_{2,i-1}} + \varepsilon_1$$
$$y_{2,i+1} = y_{2,i} + y_{1,i}\Delta t + \varepsilon_2 \tag{3.30}$$

Three sources of uncertainty - uncertain model errors (ε_1, ε_2) and the initial condition $y_{2,0}$ - are considered for GSA analysis and are propagated through the coupled system first as independent Gaussian random variables with statistics shown in Table 4.1 and then as dependent variables with the correlation structure shown in Eq. (3.31). Negative correlation between model discrepancy parameters was observed after calibration in Refs. [27] and [29], with some correlations tending toward negative 1, thus, a negative correlation ρ = -0.5 is imposed between model errors ε_1 and ε_2 in Eq. (3.31).

Table 4.1. Uncertain model inputs and errors

Parameter	Mean	Variance
$y_{2,0}$	20	4
ε_1	0	4e-4
ε_2	0	4e-4

$$\rho = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -0.5 \\ 0 & -0.5 & 1 \end{bmatrix} \tag{3.31}$$

A 3-dimensional Sobol' sequence was generated for $y_{2,0}$, ε_1, ε_2 and propagated through to model outputs y_1 and y_2 for t_f = 50*s*. Figure 4.4 demonstrates the prediction uncertainty in both y_1 and y_2 1) when just the uncertainty in the input condition is propagated as a random variable (RV) with ε_1 and ε_2 fixed at their mean values (black), 2)when all three sources of uncertainty are propagated as RVs (blue), and 3) when correlation among ε_1 and ε_2 is considered (green). It is observed in Figure 4.4a that a negative correlation between ε_1 and ε_2 slightly increases uncertainty in y_1 predictions compared to the independent case however, Figure 4.4b demonstrates that this

correlation decreases the uncertainty in y_2 predictions. This is considered to be due to the inversely proportional relationship between y_1 and y_2, but the additive relationship between y_2 and y_1. Note that since ε_1 and ε_2 both have zero mean, the mean predictions overlap for each of the three cases in Figure 4.4a and Figure 4.4b.

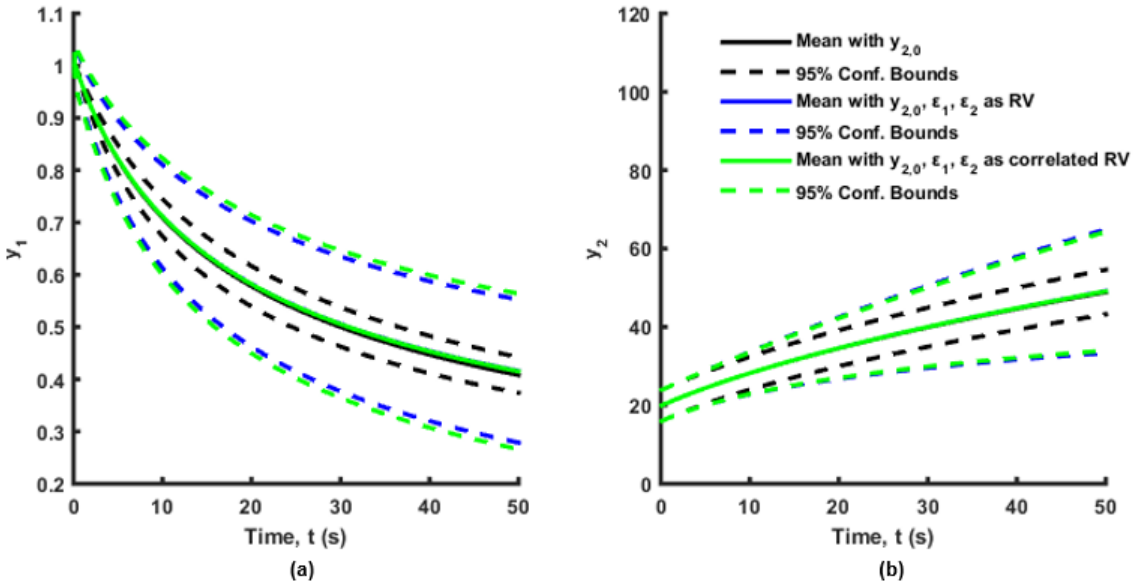

Figure 4.4. Coupled y_1 predictions (a) and y_2 predictions (b) through t = 50s

The sensitivities throughout the time to $y_{2,0}$, ε_1, and ε_2 for each model were computed using the generalized ISK-GSA methodology for both the independent and dependent parameter cases. The first-order (Eq. (3.27)) and the total effects Sobol' indices (Eq. (3.29)) were computed with a 3-dimensional Sobol' sequence of N = 1,000 which were assumed to be the conditional cumulative density function (CDF) values in each dimension d. Note that in the dependent parameter case, an iterative procedure was used to map the conditional CDF function values to the parameter space, however in the independent case the marginal CDF functions could be used.

The first-order effects and total effects for model output y_1 for the independent and correlated cases are shown in Figure 4.5. First, is observed in Figure 4.5a and Figure 4.5b that the first-order and total effects on y_1 are equivalent in the independent parameter case indicating that the

parameter interactions within the model itself are insignificant. The sensitivity of y_1 to the initial condition $y_{2,0}$ decreases over time, as expected, however, there is a trade-off between the effects ε_1 and ε_2 through time as the model output y_1 becomes more sensitive to the uncertainty in ε_2 due to model interactions.

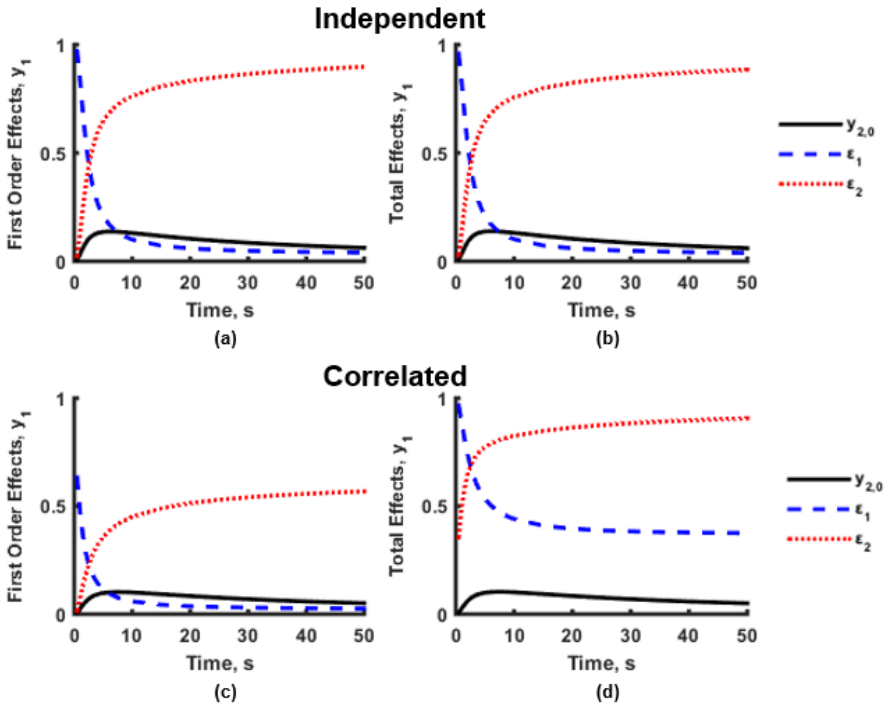

Figure 4.5. First-order and total effects on model output y_1 with independent (top) and correlated (bottom) parameters

In contrast to the independent variable case, Figure 4.5c shows that individual first-order effects of both ε_1 and ε_2 change through time when correlation is considered. First, the first order effects of ε_1 immediately decrease at t = 0s due to the correlation of -0.5 with ε_2. Then, the first order effects of ε_2 do not approach 1 as in the independent case, but rather approach approximately 0.65. However, y_1 sensitivity to $y_{2,0}$ in Figure 4.5c undergoes little change between the independent and dependent cases since $y_{2,0}$ itself is not correlated with any other parameter. For the dependent variable case in Figure 4.5d, however, the total effects of ε_2 immediately increase from 0 to 0.35 at t = 0s due to the correlation with ε_1 and it is ε_1 that approaches approximately 0.35 as opposed

to 0 in the independent case. This is intuitive since the total effects Sobol' index captures parameter interactions regardless if they are intrinsic to the model or the parameters themselves. Note that these ISK-GSA estimates maintain the summation criteria of $\Sigma S_{d,1} \leq 1$ and $\Sigma S_{d,T} \geq 1$ of the first order and total effects, respectively, over all time instants.

The first order and total effect sensitivity indices for the y_2 prediction are shown in Figure 4.6 with similar conclusions drawn from the sensitivities of y_1 from Figure 4.5. First, the first-order sensitivities in Figure 4.6a and total effect indices in Figure 4.6b are equivalent in the independent case due to negligible parameter interactions within the model. Second, y_2 sensitivity to $y_{2,0}$ undergoes little change between the independent and dependent variable cases since $y_{2,0}$ itself is not correlated with any other parameter.

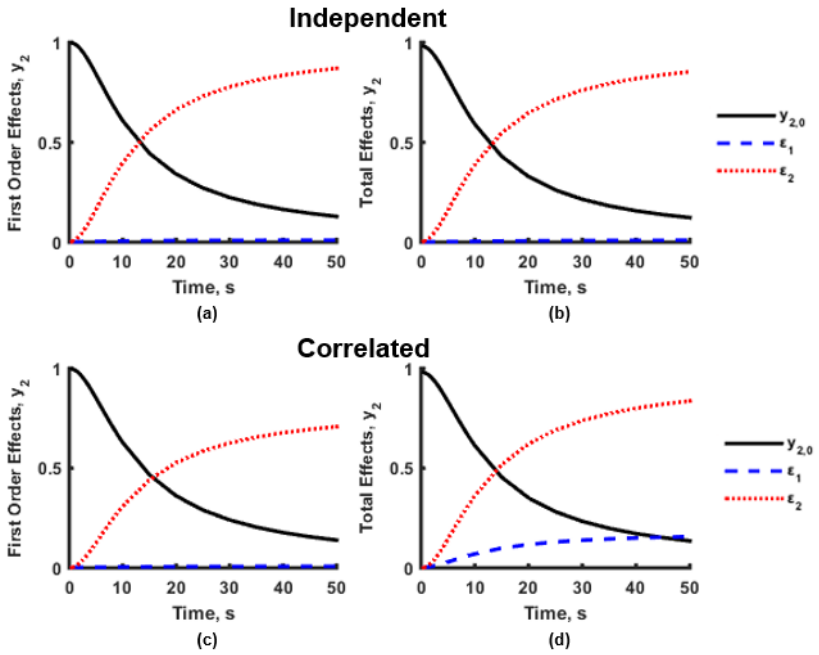

Figure 4.6. First-order and total effects on model output y_2 with independent (a-b) and correlated (c-d) parameters

For the dependent variable cases in Figure 4.6c-d, it is observed the first-order effects of y_2 in Figure 4.6c sum close to unity since the sensitivity of y_2 to ε_1 is low, however, in Figure 4.6d the

total effects sensitivities of y_2 to both ε_1 and ε_2 increases over time due to the effect of correlation between ε_1 and ε_2.

Figure 4.7a-b illustrate differences between positive and negative correlations on the first order and total effect sensitivities of both model outputs to error parameter ε_2 through time. First, it is observed that as the correlation increases from -0.5 to -0.9, the first order effects of ε_2 on y_1 become closer to 0. It can be shown further that as the correlation approaches -1, the first order effects of both ε_1 and ε_2 approach 0, whereas the total effects in Figure 4.7b approach 1.

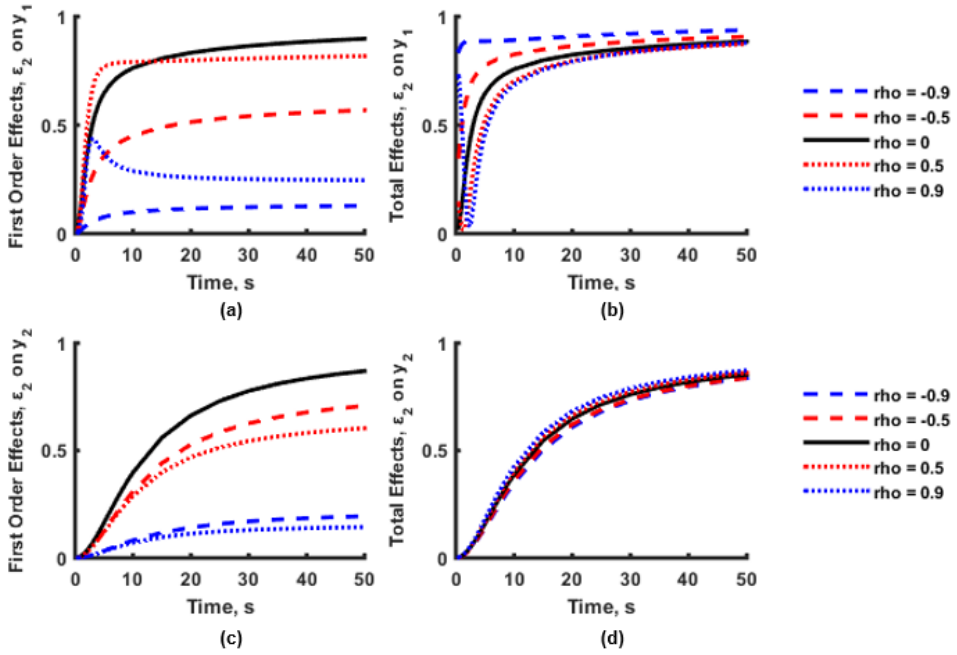

Figure 4.7. Effect of ε_1 and ε_2 correlation on the sensitivities of y_1 to ε_2 (a-b) and y_2 to ε_2 (c-d)

However, when the correlation between the parameters is positive, it is observed in Figure 4.7a-b that the first-order effects of ε_2 on y_1 are greater than the total effects. Again, this is attributed to the inversely proportional relationship between $y_{1,i}$ and $y_{2,i\text{-}1}$ (and, by extension, ε_2), and this phenomenon does not occur in the first order and total effects of ε_2 on model output y_2 in Figure 4.7c-d. In both positive and negative correlation cases in Figure 4.7c-d, the first order effects of ε_2

on model output y_2 decreased compared to the independent ($\rho = 0$) case, with the positive correlation cases exhibiting lower first order effects over time and higher total effects than the negative correlation cases. It is interesting to note that the independent case in Figure 4.7c represents the maximum first order effect, however the $\rho = 0$ case does not bound the total effect sensitivities in Figure 4.7d. In general, the same is true between the first order and total effects on y_1 in Figure 4.7a-b, however the relationship between the positive and negative correlations with the independent case in the total effects are opposite, with the positive correlations having lower total effects than the independent case for y_1 but higher total effects for y_2.

For this 3-parameter example ($D = 3$), the sensitivity convergence rates between the independent and correlated cases were equal using the generalized ISK-GSA methodology with a Sobol' sequence. Here, the first-order indices of both y_1 and y_2 on average converged in 250 iterations using a 1-D Sobol' sequence. However, the total effect indices converged in approximately 800 iterations from computing conditional expectation estimates in D-1 dimensions. The comparison of ISK-GSA convergence rates with a Sobol' sequence for a four dimensional aerothermal application problem ($D = 3$) will occur next in Section 4.7 as well computing post-calibration sensitivities of the aerothermal models using the generalized ISK-GSA methodology and existing model runs from calibration.

4.7 Application Example: Coupled Aerothermal Sensitivities

A global sensitivity analysis is performed on coupled aerothermal models to quantify the sensitivity of previously calibrated coupled model errors and uncertainty on aerodynamic heating and heat transfer predictions through time.

4.7.1 Aerothermal Model Calibration

Building on the Bayesian model calibration work in Chapters 2 and 3, the aerodynamic heat flux Q predicted by Eckert's reference temperature method [33] is coupled with the one-dimensional heat balance equation to predict the structural temperature $T_{structure}$ and wall temperature T_w of a panel on a hypersonic vehicle upstream of an oblique shock.

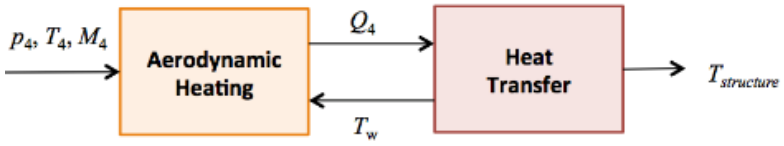

Figure 4.8. Aerothermal coupling

Aerodynamic heating errors $\delta_{Q,t}$ and heat transfer errors $\delta_{T,t}$ propagate through each iteration and are functions of the changing model inputs according to Eqs. (3.32) and (3.33). The sources of uncertainty considered in the post-calibration sensitivity analysis are the four model error parameters $\boldsymbol{X}$ = [c_0, c_1, d_0, d_1].

$$\delta_{Q,t}(T_{w,t-1}) = c_0 + c_1 \frac{T_{w,t-1} - T_{w,0}}{T_4} \tag{3.32}$$

$$\delta_{T,t}(Q_{4,t}) = (d_0 + d_1 Q_{4,t})\Delta t \tag{3.33}$$

Based on aerothermal tests conducted in the NASA Langley HTT tunnel [20], the reported heat flux measurements at t = 0 were used to construct a 3 second temperature history at a rates of Δt = 0.05s (20 samples per second). Measurement noise was assumed to be zero-mean Gaussian with a standard deviation of 1K. This temperature history was used to calibrate the coupled heat flux and temperature predictions through time and the prior and posterior predictions with 95% confidence bounds are shown in Figure 4.9a and 4.9b, respectively. The calibration was completed using 10^3 slice samples [54] and the post-calibration prediction confidence in temperature increased across

all time steps by reducing uncertainty in both Eckert's reference temperature and 1-D heat transfer predictions.

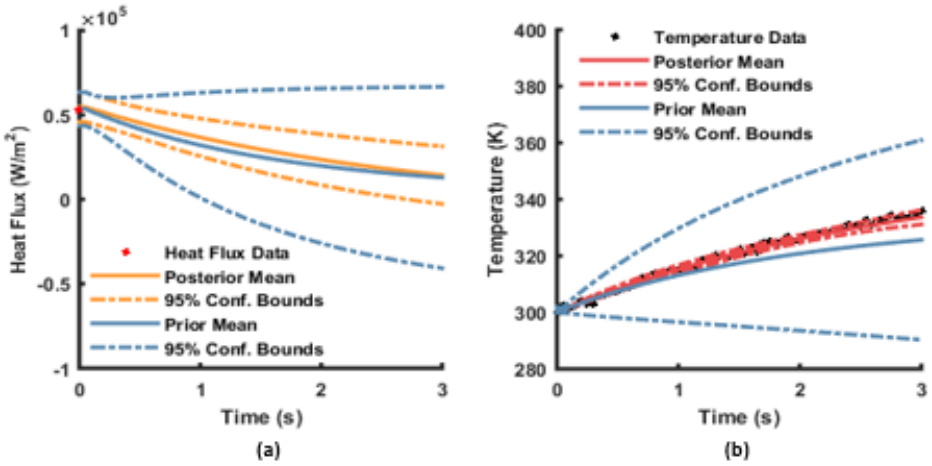

Figure 4.9. Prior and posterior heat flux (a) and temperature (b) predictions

The posterior correlation among the four discrepancy parameters are shown in Eq. (3.34) and provides one application for the generalized ISK-GSA methodology for dependent variables. In addition, both the prior and posterior model discrepancy distributions of δ_Q and δ_T are correlated as well, and it is in this space that we apply the generalized ISK-GSA methodology to the coupled, aerothermal analysis. The prior and posterior correlations among the model discrepancy terms δ_Q and δ_T are depicted in Figure 4.10a-b first at the beginning of the analysis at t_0 = 0s and then at the end at t_f = 3s.

$$\rho_X = \begin{bmatrix} 1 & -0.29 & -0.42 & 0.02 \\ -0.29 & 1 & -0.54 & 0.10 \\ -0.42 & -0.54 & 1 & -0.59 \\ 0.02 & 0.10 & -0.59 & 1 \end{bmatrix} \tag{3.34}$$

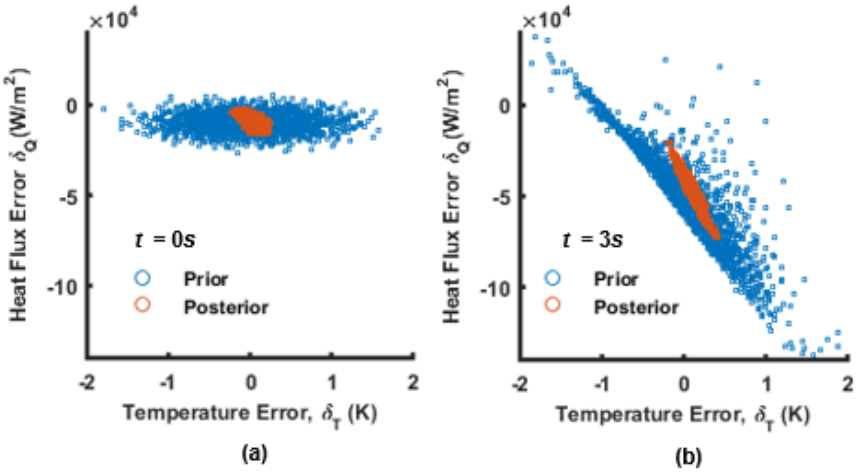

Figure 4.10. Joint prior and posterior distributions for δ_Q and δ_T at (a) t = 0s and (b) t = 3s

At t = 0s in Figure 4.10a, correlation ρ_δ between model discrepancies is observed to be -0.6 among the posterior samples wheras ρ_δ = 0 among the prior model discrepancy predictions since the discrepancy parameters $\boldsymbol{\theta}$ were independent before calibration. However, strong negative correlations are present among both prior model discrepancy predictions at t = 3s in Figure 4.10b. First, Figure 4.10b shows the prior correlation between δ_Q and δ_T becomes -0.9 solely due to discrepancy interactions that are inherent to the prediction.. In comparison, the posterior correlation between δ_Q and δ_T changes from -0.6 at t = 0s and grows to -0.95 at t =3s in Figure 4.10b from model interactions through time. The trend of both the prior and posterior model error correlations through time in Figure 4.11 demonstrates that the input parameter correlations affect the coupled prediction until the coupling effects become more prominent over time, and extending the analysis through demonstrates the convergence of both prior and posterior error correlations to -0.9.

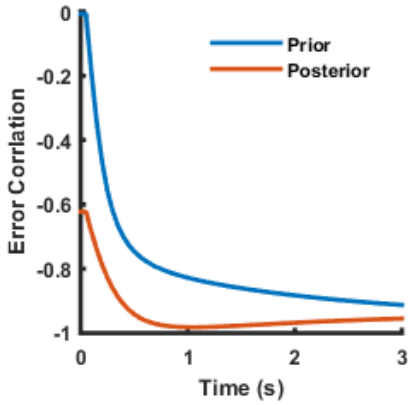

Figure 4.11. Prior and posterior correlation between δ_Q and δ_T

Note that the generalized ISK-GSA methodology is also needed to compute the prior sensitivities of the model outputs to δ_Q and δ_T through time as well, since the errors become correlated through time due to the model interactions. The difference however, lies in how the input-output samples are generated, where prior sensitivity analysis uses 4-D Sobol' sequence and the posterior analysis uses samples generated from the MCMC algorithm.

4.7.2 Pre-Calibration Aerothermal Sensitivities Through Time using Generalized ISK-GSA

The pre-calibration sensitivity estimates of heat flux and temperature to model discrepancies δ_Q and δ_T were performed using generalized ISK-GSA with Sobol' sequences. Figure 4.12 shows the sensitivity convergences observed during a prior sensitivity analysis using the generalized ISK-GSA method with Sobol' sequences compared to latin-hypercube design of the input space.

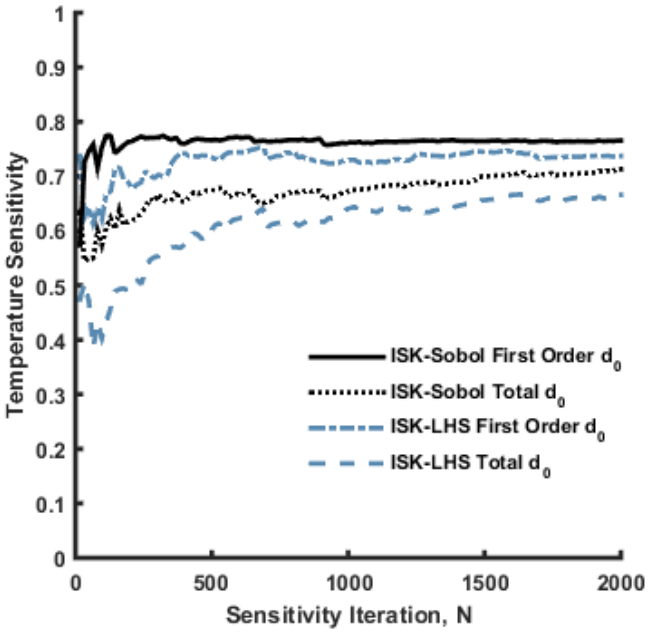

Figure 4.12. Convergence of prior sensitivities of temperature to d_0 at t = 3s using latin-hypercube and Sobol' sequences

Similar to the examples in Section 4.4, the first-order effects converged using fewer model evaluations than the total effects. Furthermore, the Sobol' sequence of the input space led to faster convergence of the prior first-order sensitivities, where the first-order effects required approximately 600 iterations to satisfy the convergence criteria (<0.1% difference in sensitivities between tests occurring at every 10^{th} iteration) compared to the 1300 iterations requires by the LHS design. It was determined that neither the total effects from the Sobol' index nor the LHS design meet the convergence criteria within in 2000 iterations for this 4-dimensional problem.

4.7.3 Post-Calibration Aerothermal Sensitivities Through Time using Generalized ISK-GSA

The post-calibration sensitivity analysis for the aerothermal example was performed using the last 4e3 posterior samples from the 10e3 slice-sampling from calibration to reduce the computational effort required to compute the ISK-GSA results at each time instant. The posterior heat flux sensitivities through time in Figure 4.13 demonstrate that the negative correlation between δ_Q and δ_T at all time instants first decreases the first order effect sensitivities of δ_Q from

1 (observed in the prior case) to 0.6, which corresponds to the negative posterior discrepancy correlation of -0.6 seen at t = 0s in Figure 4.13. Then, the posterior sensitivity of Q_n to model discrepancy δ_Q follows a similar trend to the posterior discrepancy correlation in Figure 4.13. This is because a correlation close to -1 indicates strong interactive effects, which in turn reduce the first order effects of the variable. Similarly, the posterior total effect of δ_T on heat flux have increased to 0.4 at t = 0s due to the correlation with δ_Q, approach 1, and then asymptotically approach approximately 0.85, which can be considered as the total effect sensitivity from interactions inherent to the model.

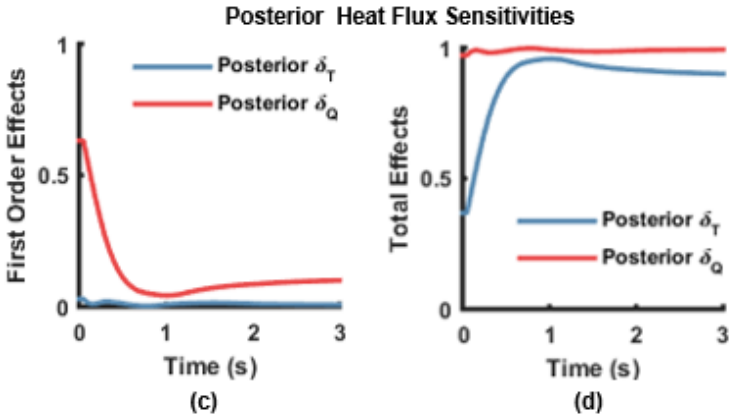

Figure 4.13. Posterior heat flux sensitivities to δ_Q and δ_T through time

The posterior sensitivities of the predicted temperature to the model discrepancies δ_Q and δ_T through time are shown in Figure 4.14. First, in contrast to the heat flux sensitivities in Figure 4.13, the first order effects in Figure 4.14a indicate influence from a source of positive correlation because, rather than the first order effects of δ_T decreasing from treating the correlation as interactions, the first order effects of δ_Q increase and both first order effects sum to greater than 1 at t = 0s. Further, the total effects in Figure 4.14a sum to less than 1, and in this case, temperature sensitivities to δ_T decrease. This prompts further investigation into the relationships between temperature T and δ_Q and δ_T through time and the effect of positive correlations on GSA sensitivity indices of coupled model outputs through time.

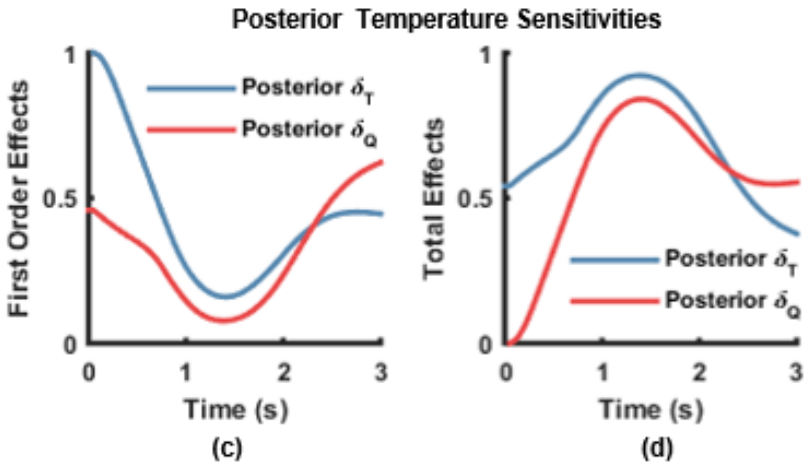

Figure 4.14. Posterior temperature sensitivities to δ_Q and δ_T through time

4.8 Conclusion

Global sensitivity analysis computations for independent variables were improved first by pairing quasi-random number generators with the ISK-GSA methodology. Furthermore, efficient GSA computations were extended to correlated variables by generalizing the ISK-GSA method and applying it to coupled models interacting through time and then to existing posterior input-output samples from Bayesian calibration. A summary of conclusions drawn from this chapter are as follows:

1. When paired with quasi-random sequences, the ISK-GSA sensitivity estimate asymptotically converges to the true sensitivity. Therefore, a sensitivity convergence criterion can be used to minimize the number of model evaluations needed for global sensitivity analysis among independent variables (e.g., before calibration).
2. The generalized ISK-GSA methodology for dependent variables improved the efficiency of GSA for:
    a) coupled model predictions, where parameters interactions within the models become more prominent over time; and

b) post-calibration sensitivity analyses by allowing the direct use of correlated posterior samples obtained from calibration.

3. The effects of both positive and negative correlations on both the first order and total effects were demonstrated. For the simplified time-dependent example, negative correlations between model errors maintained the expected summation criteria for first-order and total effects, however, positive correlations did not. Similar results were observed for the aerothermal problem, and thus, the effects of positive correlations on the sensitivities of coupled models require further understanding.
4. The generalized ISK-GSA method led to a broader analysis of sensitivity in coupled, time-dependent analyses where first-order and total effects are influenced by both parameter correlation and model interactions through time. For example, in the aerothermal application example, correlation effects on the sensitivity estimates were seen in early time instances until coupling effects became more significant through time.

In future work, this methodology may be embedded directly into MCMC sampling procedures to monitor the evolution of the posterior parameters after each sample. In this way, sources of uncertainty can be down selected adaptively as the likelihood is explored in the MCMC algorithm. This may also account for the possibility that some regions of the likelihood may be more sensitive to a subset of the parameter space than other regions. Also, at a given point in the calibration the generalized ISK-GSA sensitivity result may indicate convergence globally or within the local likelihood region, which may inform the optimal step-sizes taken in each dimension and further improve calibration convergence.

CHAPTER 5

MODEL SELECTION AND COUPLING IN TIME-DEPENDENT, MULTIDISCIPLINARY SIMULATIONS

5.1 Introduction

The focus of this dissertation has been on multidisciplinary analyses where several single-discipline models are linked for prediction. In such time-dependent, multidisciplinary simulations, the following analysis decisions are made *a priori* that affect both the prediction accuracy and simulation cost: 1) the fidelity of each model component in the simulation hierarchy where reduced-order, reduced-physics, or low-fidelity models may be used in place of higher-fidelity models, and 2) the characteristic time step for each disciplinary model component where the time step ratios reflect the coupling between disciplines in time-dependent analyses. For example, in transient aerothermoelastic simulations, the model fidelities and time steps maybe chosen to capture the response at critical points along the trajectory (e.g., high-risk maneuvers, takeoff and landing). However, these *a priori* model fidelity and time step choices may be unnecessarily fine during the portions of the trajectory with slower-moving phenomena (e.g., cruise, elevation changes), adding unneeded cost to the analysis. In the case of a coarse model or time step being selected due to simulation cost restrictions, prediction errors may be introduced in the analysis. Thus, adaptive model fidelity and time step selection methods are needed for coupled, time-dependent analyses to balance both prediction cost and accuracy during long duration simulations.

In traditional statistical analysis, model selection criteria are based on two factors: a) the likelihood the model explains the data (accuracy) and b) the number of predictors (complexity – the smaller the better) [84]. Model selection methods are well-developed for choosing among candidate statistical models [85], reduced-order models and approximations for single-disciplinary

analysis, and spatial meshing and scale resolution in multi-scale problems [86]. This chapter expands these ideas to time-dependent, coupled models where complexity and goodness-of-fit characteristics are less explicitly defined.

In the context of UQ, however, uncertainty propagates through time and through the models and their coupling interfaces in the forward prediction. Thus, there is a need for a selection metric that uses all the available probability information rather than an average or single summary statistic in order to comprehensively and robustly assess the prediction confidence and inform the best adaptive procedure. The model reliability metric developed in model validation research helps quantify the confidence by comparing stochastic predictions to experimental data or high-fidelity simulations [66]. The use of model reliability as a model selection criterion was developed by Hombal and Mahadevan [87].

The methods developed in this chapter expand on work by Hombal and Mahadevan by considering time-dependent, multidisciplinary analyses where the goal is to maintain a desired level of prediction accuracy through time under a limited computational budget. The methods are applied to the time-dependent aerothermal models that were calibrated in Chapter 3 and studied using ISK-GSA for sensitivity analysis in Chapter 4.

Two prediction performance measures – prediction accuracy and prediction reliability – are presented in Section 5.2.1. Subsequently, the accuracy- and reliability-based loss functions are proposed in Section 5.3 and resulting optimizations with each over both time step and coupling from are demonstrated an illustrative example. Section 5.4 then demonstrates the optimization of the coupled aerothermal simulation through time using the reliability-based loss function.

5.2 Simulation Accuracy, Reliability, and Cost

Consider a coupled, time-dependent analysis with two disciplinary models in which the load x interacts with the response of the system y. Two single-disciplinary model components $\mathcal{M}_x$ and $\mathcal{M}_y$ are used to simulate the interactions between x_n and y_n at each time step Δt_n for $n = 1$ to N according to Figure 5.1. The initial condition y_0 at initial time $t_0 = 0$ seconds is considered known and N is the number of time steps between t_0 and the final time of interest $t_N = t_f$. Assuming the cost of the coupled analysis at each time instant is C_n, the total simulation cost C_f is computed in Eq. (4.1) as the sum of the costs incurred after each time step.

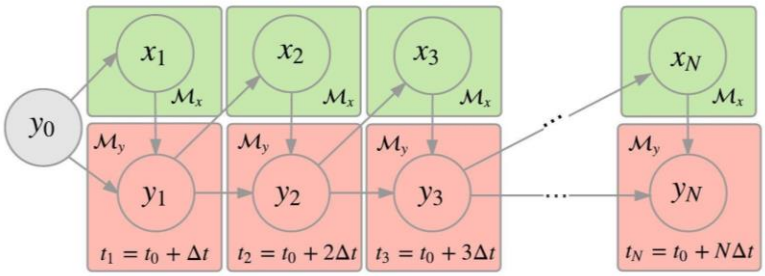

Figure 5.1. Partitioned simulation with fixed coupling and time step

$$C_f = \sum_{n=1}^{N} C_n = \sum_{n=1}^{N} \left[C_{x,n} + C_{y,n} \right] \tag{4.1}$$

Suppose that there exists set of single-disciplinary modeling alternatives $\mathcal{M}_x'$ and $\mathcal{M}_x'$ that may be of higher or lower fidelity than $\mathcal{M}_x$ and $\mathcal{M}_y$. Figure 5.2 illustrates a partitioned analysis that uses these alternative analyses as well as a zero-cost analysis $\emptyset_x$ that assumes the effect of y on x is negligible at time t_n (i.e., $x_n = x_{n-1}$). Therefore, the simulation cost at the n^{th} time step in Eq. (4.1) is further decomposed into the cost of the n^{th} model fidelity combination $F_n = \{\mathcal{M}_x^*, \mathcal{M}_y^*\}$ where each single-disciplinary model (i.e., $\mathcal{M}_x^* \in \{\mathcal{M}_x, \mathcal{M}_x', \emptyset_x\}$ and $\mathcal{M}_y^* \in \{\mathcal{M}_y, \mathcal{M}_y', \emptyset_y\}$) is associated with a known model cost (i.e., $C_x^* \in \{C_x, C_x', 0\}$ and $C_y^* \in \{C_y, C_y', 0\}$), respectively.

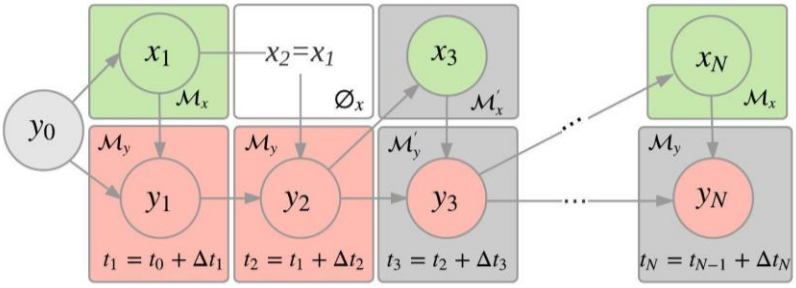

Figure 5.2. Partitioned simulation with variable coupling and time step

Figure 5.2 also demonstrates the use of different time steps sizes Δt_n^* for each combination in $F_n = \{\mathcal{M}_x^*, \mathcal{M}_y^*\}$. Intuitively, increasing the time step size will reduce the number of time integrations needed for fixed-duration analyses and would subsequently reduce the overall simulation cost. However, increasing the time step size may have adverse effects on the prediction accuracy given the choice of models in F_n. Thus, the goal of this chapter is to develop a decision-making framework for the coupling and temporal fidelity selection of $F_n = \{\mathcal{M}_x^*, \mathcal{M}_y^*, \Delta t_n^*\}$ at each time step n that balances both simulation cost and prediction accuracy. In Section 5.2.1, two metrics for prediction accuracy – one deterministic (error) and one stochastic (model reliability)- are discussed.

5.2.1 Prediction Accuracy and the Model Reliability Metric

Consider a set of data realizations D_n at time t_n from a known probability distribution $p_{Dn}(d_n)$. When comparing against a deterministic prediction y_n, Eq. (4.2) defines the expected prediction error $E[\varepsilon_n]$ as the magnitude of the difference between prediction y_n and expected value of the data $E[D_n]$. Prediction accuracy is assessed deterministically by comparing the expected prediction error $E[\varepsilon_n]$ against a prediction error tolerance $\tau_n \geq 0$.

$$E[\varepsilon_n] = \left| y_n - E[D_n] \right| = E\left[\left| y_n - D_n \right| \right] \tag{4.2}$$

However, several sources of uncertainty may contribute to uncertainty about the prediction y_n (e.g., natural input variability, model form error, etc.) so that y_n may no longer be considered a deterministic quantity but be characterized by a probability distribution $p_{Yn}(y_n)$.Thus, both data D_n and the model prediction y_n are stochastic; the second equality in Eq. (4.2) applies in this case. The prediction error variance $Var[\varepsilon_n]$ is the sum of prediction and data variances in Eq. (4.3).

$$Var[\varepsilon_n] = Var[Y_n] + Var[D_n] \tag{4.3}$$

Errors and uncertainty result from imperfect knowledge and numerical errors and are of two types - model form uncertainty and solution approximations (e.g., spatio-temporal discritization) [26] – that aggregate at each stage of the analysis based on the chosen models and coupling characteristics (e.g., monolithic vs. partitioned, high-fidelity vs. low-fidelity, strong vs. weak coupling). The model reliability metric in Eq. (4.4) assesses the agreement between distributions $p_{Yn}(y_n)$ and $p_{Dn}(d_n)$ on a scale from 0 and 1. The model reliability metric can be calculated for any distributions of Y_n and D_n using Monte carlo sampling; the first equality in Eq. (4.4) represents this general case. The second equality in Eq. (4.4) demonstrates the model reliability following Gaussian assumptions for $p_{Yn}(y_n)$ and $p_{Dn}(d_n)$, where standard normal cumulative distribution function (CDF) denoted by Φ is used.

$$R_n = \Pr\left[\left|y_n - D_n\right| \le \tau_n\right] = \Phi\left[\frac{\tau_n - E\left[\varepsilon_n\right]}{\sqrt{Var\left[\varepsilon_n\right]}}\right] \tag{4.4}$$

The model reliability metric inherently includes the prediction accuracy by comparing the prediction error $E[\varepsilon_n]$ against the error tolerance τ_n at time t_n, but it is additionally influenced by the uncertainty that propagates through the simulation to y_n (Eq. (4.3)). Therefore, a target prediction reliability R_{lim} may not be achieved even if $\tau_n - E[\varepsilon_n] \ge 0$ (i.e., the mean prediction error

is less than the error tolerance) due large prediction error variance $Var[\varepsilon_n]$ from either prediction uncertainty or data uncertainty. These accuracy and precision tradeoffs in the prediction reliability metric are explored using an illustrative example problem in Section 5.2.2 which is used further to develop the adaptive model selection methodology in Section 5.3.

5.2.2 Illustrative Example

Two models, $\mathcal{M}_x$ in Eq. (4.5) and $\mathcal{M}_y$ in Eq. (4.6), are representative of multidisciplinary, time dependent analysis. The models are coupled such that x_n is the rate of change between y_n and y_{n+1} across the time step Δt_n and x_n itself is the ratio between the current prediction y_n and the initial condition y_0. The initial condition y_0 is uncertain and characterized by a normal distribution $p_{Y0}(y_0) \sim \mathcal{N}(20,2)$ from which 1000 samples are randomly drawn and propagated through the coupled analysis. Therefore, the mean prediction and 95% confidence bounds of x and y using a model fidelity and time step combination of $F = [\mathcal{M}_x, \mathcal{M}_y, \Delta t = 1s]$ are shown in Figure 5.3a-b.

$$x_n = \mathcal{M}_x = \frac{y_0}{y_{n-1}} \tag{4.5}$$

$$y_n = \mathcal{M}_y = y_{n-1} + x_n \Delta t_n \tag{4.6}$$

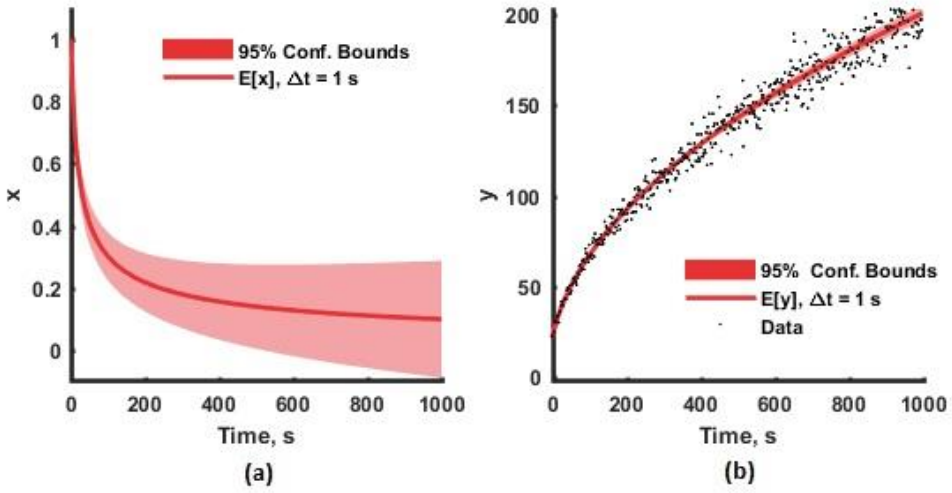

Figure 5.3. Coupled simulation $F = [\mathcal{M}_x, \mathcal{M}_y, \Delta t = 1s]$ compared to data D_n

Random realizations of data D_n (one of which is shown in Figure 5.3b) were generated using a non-stationary Gaussian measurement noise of 5% of the true output y_{true} and an observation rate of $\Delta t_D = 0.1$s. The true quantities x_{true} and y_{true} were simulated from the initial condition $y_{0,true} = 20$ using model fidelity and time step combination $F = [\mathcal{M}_x, \mathcal{M}_y, \Delta t = 0.001s]$ through $t_f = 1000s$. Recall that prediction accuracy is computed as the difference between the prediction error $E[\varepsilon_n]$ in Eq. (4.2) and an error tolerance $\tau_n \geq 0$. In Figure 5.4, the error tolerance is a percentage – either 1%, 2%, or 5% – of the magnitude of the expected value of the data $E[D_n]$ at time t_n.

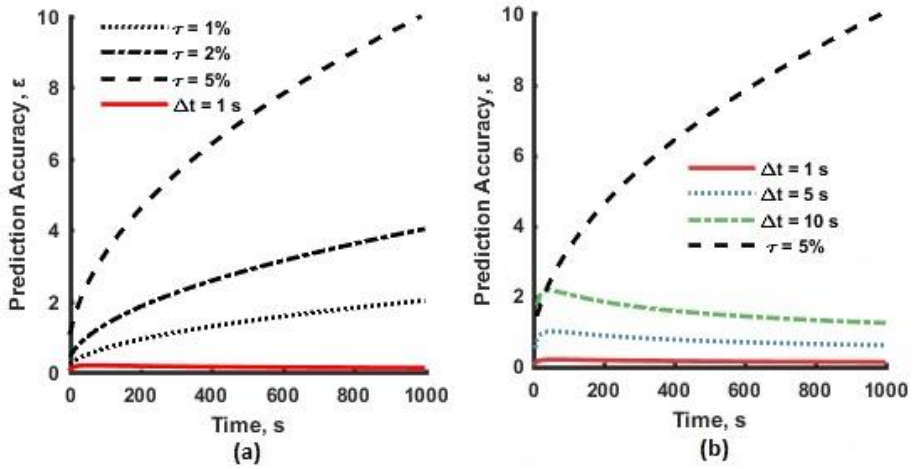

Figure 5.4. Prediction error ε_n with (a) $\Delta t = 1s$, $\tau = 1\%$, 2%, and 5% and (b) $\tau = 5\%$, $\Delta t = 1$, 5, and 10s

Figure 5.4a shows that the prediction error observed throughout a fixed time step analysis (Δt = 1s) does not exceed even the lowest error tolerance limit of τ = 1%. Figure 5.4b then demonstrates that increasing the simulation time step increases the prediction error, where an analysis with a fixed time step of Δt = 10s exceeded the error limit around t = 40s. These conclusions are offered the within the context of the model reliability metric in Figure 5.5 so that comparisons can be drawn between the two methods.

The results in Figure 5.5a demonstrate the role that the choice of error tolerance τ has on the model reliability metric where for a fixed time step analysis (Δt = 1s) the prediction reliability is close to 1 when the error tolerance is large (τ = 5%) and near to 0.6 when the tolerance is more restrictive (τ = 1%). In Figure 5.5b, a large simulation time-step of Δt = 10s with a large error tolerance of τ = 5% is shown to have a reliability of 0.4 at early time instances, however, meet the desired reliability R_{lim} of 0.9 during a large portion of the analysis.

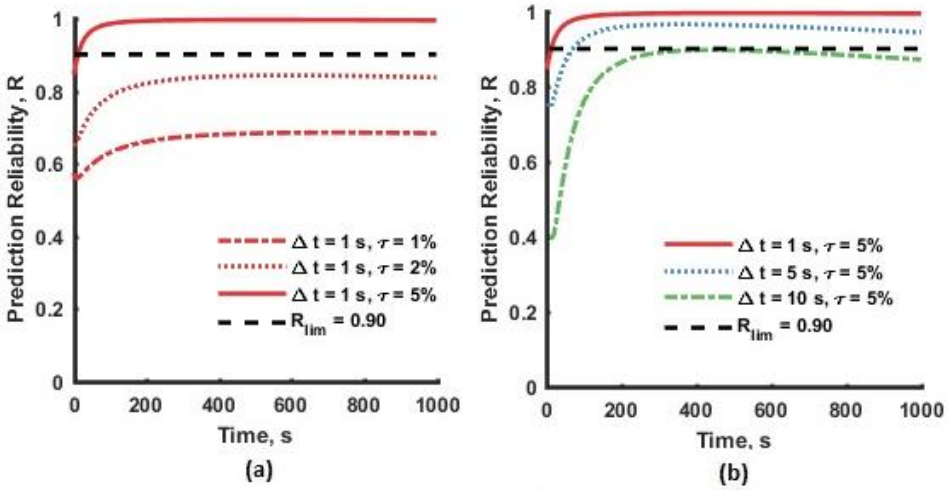

Figure 5.5. Prediction reliability R_n with (a) Δt = 1s, τ= 1%, 2%, and 5% and (b) τ = 5%, Δt = 1, 5, and 10s

Table 5.1. Simulation statistics, reliabilities, and costs at t_f=1000*s* with τ = 5% for Δt = 1, 5, and 10*s*

	$\sigma_{x,f}$	$\sigma_{y,f}$	$E[\varepsilon_f]$	R_f	N	C_f
Δt = 1s	9.59e-2	2.21	0.11	0.99	1000	4000*u*
Δt = 5s	9.61e-2	4.95	0.59	0.94	200	800*u*
Δt = 10s	9.64e-2	7.05	1.22	0.87	40	400*u*

Table 5.1 further demonstrates that an analysis that uses a fixed time step of Δt = 10s would result in 90% cost savings (compared to Δt = 1s) at a 3% deficit of meeting the target reliability R_{lim} of 0.9 at t_f = 1000s. Therefore, Section 5.3 develops a bi-objective optimization methodology to balance cost vs. error and cost. vs reliability objectives.

5.3 Proposed Coupling and Temporal Fidelity Selection Methodology

A flowchart for the adaptive model selection procedure to select F_{n+1}^* based on a loss function L at time t_{n+1}^* for the two-discipline system is summarized in the flow chart in Figure 5.6. First, the limiting values of each quantity of interest τ_n, R_{lim}, C_{lim} are determined by an analyst.

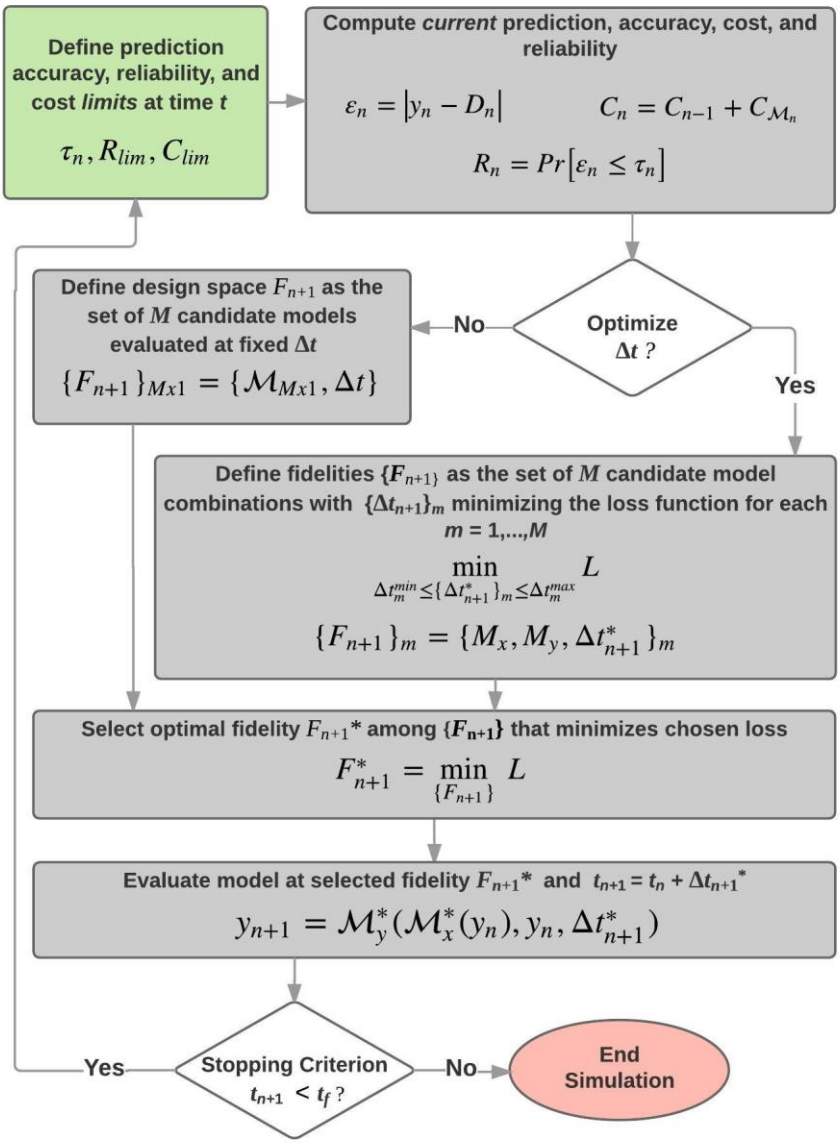

Figure 5.6. Flowchart of the process of finding optimal coupling and temporal fidelity at time t_{n+1}

Two prediction performance measures – prediction accuracy and prediction reliability – were presented in Section 5.2.1 and the proposed accuracy- and reliability-based loss functions used to optimize coupling and temporal fidelities are presented in Section 5.3.1. Sections 5.3.2 and 5.3.3 compare the two loss functions on an illustrative example and 5.3.4 demonstrates the reliability-based loss function on combined coupling and temporal fidelity selection.

5.3.1 Accuracy- and Reliability-Based Loss Functions

The proposed accuracy-based loss function L_ε shown in Eq. (4.7) is minimized to select the optimal coupling and temporal fidelity combination $F_{n+1}^* = [\mathcal{M}_x^*, \mathcal{M}_y^*, \Delta t^*]_{n+1}$ among multiple candidate fidelity options. Here, the * notation indicates the parameters that are dependent on the model fidelity choice of F_{n+1}^* (including Δt_{n+1}^*).

$$L_\varepsilon(F_{n+1}^*) = \frac{w_\varepsilon}{w_\varepsilon + w_C}\left(\frac{\varepsilon_{n+1}^*}{\tau_{n+1}^*} - 1\right)^2 + \frac{w_C}{w_\varepsilon + w_C}\left(\frac{C_{n+1}^* / C_{lim}}{(t_n + \Delta t_{n+1}^*)/t_f} - 1\right)^2 \tag{4.7}$$

$$w_\varepsilon = \frac{\varepsilon_n}{\tau_n}, w_C = \frac{C_n / C_{lim}}{t_n / t_f} \tag{4.8}$$

For a candidate model coupling fidelity and time step to be considered for optimality in Eq. (4.7), the prediction error ε_{n+1}^*, error limit τ_{n+1}^*, and the cost C_{n+1}^* that would accumulate across the candidate time step Δt_{n+1}^* if selected are either known given the model or estimated. The proposed approach uses the model error and time step history obtained during the simulation to estimate $\varepsilon_{n+1}^*(\Delta t_{n+1}^* \mid \mathcal{M}_x^*, \mathcal{M}_y^*)$, considers the model costs C_{n+1}^* as known quantities, linearly extrapolates to τ_{n+1}^* from τ_n and τ_{n-1}^*.

Note that the squared components of the loss function L_ε are scaled by normalized positive weights w_ε and w_C presented in Eq. (4.8). The weight w_ε is the ratio of the current prediction error ε_n, and current error tolerance τ_n,. When the ratio $w_\varepsilon < 1$, it implies that $\varepsilon_n < \tau_n$. The weight w_C is composed of two ratios: a) the ratio of the current cost C_n to the overall cost budget C_{lim} and b) the ratio of the current simulation time t_n to the final time of interest t_f.. When the ratio $w_C = 1$, it indicates that the simulation will conclude at $t_N = t_f$ and $C_N = C_{lim}$, meaning the simulation budget is entirely exhausted.

Similar to Eqs. (4.7) and (4.8), the proposed reliability-based loss function L_R is shown in Eq. (4.9) with weights w_R and w_C shown in Eq. (4.10). The weight w_R is the ratio between the current reliability R_n and R_{lim} where R_n is placed in the denominator to prevent the current prediction reliability from tending toward zero.

$$L_R(F_{n+1}^*) = \frac{w_R}{w_R + w_C}\left(\frac{R_{lim}}{R_{n+1}^*} - 1\right)^2 + \frac{w_C}{w_R + w_C}\left(\frac{C_{n+1}^* / C_{lim}}{(t_n + \Delta t_{n+1}^*)/t_f} - 1\right)^2 \tag{4.9}$$

$$w_R = \frac{R_{lim}}{R_n}, w_C = \frac{C_n / C_{lim}}{t_n / t_f} \tag{4.10}$$

First the accuracy- and reliability-based loss functions are implemented on the illustrative example from Section 5.2.2 for optimizing two simulation cases: 1) coupling fidelity at a fixed time step in Section 5.3.2 and 2) temporal fidelity selection at fixed coupling in Section 5.3.3. Then in Section 5.3.4 the reliability-based loss function is applied for combined coupling and temporal fidelity selection.

5.3.2 Case 1: Coupling Fidelity Selection (Fixed Time Step)

Continuing with the illustrative example from Section 5.2.2, two competing model coupling options were identified, a two-physics model $\{\mathcal{M}_x,\mathcal{M}_y\}$ with cost $C_n = 4u$ and a single-physics model $\{\emptyset_x,\mathcal{M}_y\}$ with cost $C_n = 1u$. Thus, for a fixed time step analysis $\Delta t = 1s$ under a restricted computational budget of $C_{lim} = 1250u$, Figure 5.7 compares the model coupling selections made from both the L_ε-optimized and L_R-optimized simulations.

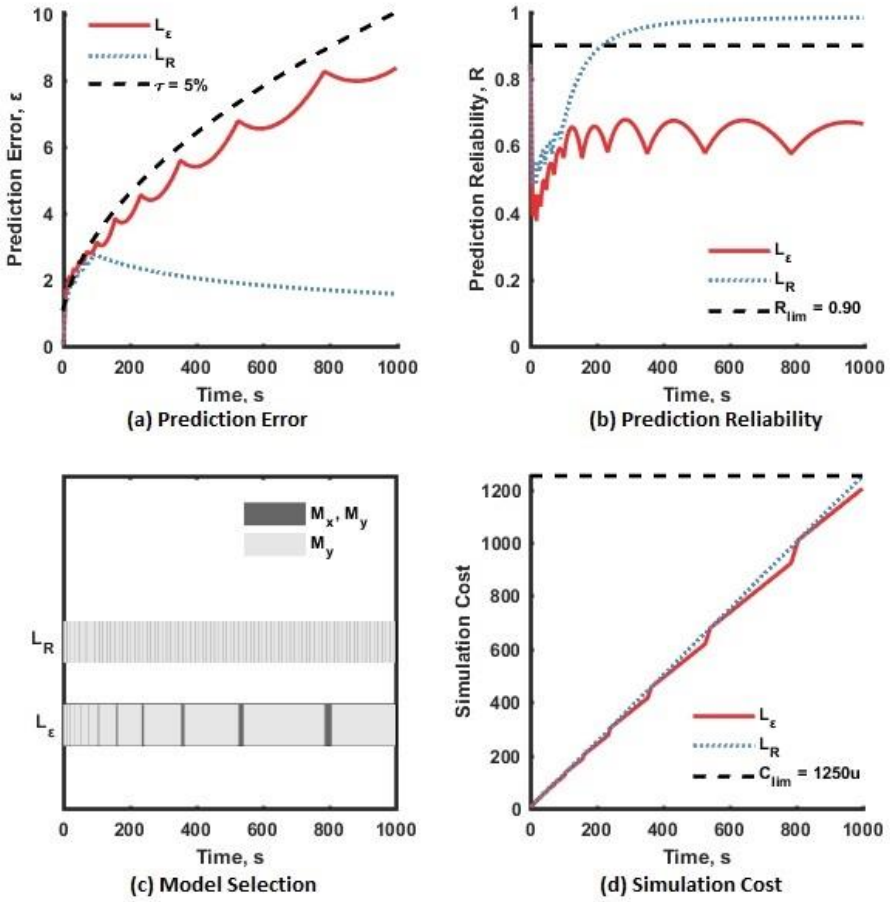

Figure 5.7. Comparison of L_ε and L_R for fixed Δt = 1s, [τ = 5%, C_{lim} = 1250u, and R_{lim} = 0.9]

Figure 5.7a-b first demonstrates that as the L_ε -optimized simulation approaches the error limit τ from using the single-physics model (which would result in an effective prediction reliability of 0.5), the L_ε-optimization then selects the two-physics model to reduce the error leading to significant jumps in simulation cost. In contrast, after the L_R-optimized simulation satisfies the reliability limit objective R_{lim} of 0.9, the cost objective is dictating the intermittent selection of the two-physics model such that the computational budget C_{lim}=1250 is exhausted completely at the end of the analysis.

5.3.3 Case 2: Temporal Fidelity Selection (Fixed Coupling))

Next, consider the simulation in Figure 5.8 where the temporal fidelity Δt_n is optimized for a fixed coupling case with the two-physics model $\{\mathcal{M}_x, \mathcal{M}_y\}$ with a cost of C_n = 4u. Here, a

computational budget $C_{lim} = 400u$ was imposed, meaning only 100 time instances can be evaluated across the 1000 second time history.

From Table 5.1 in Section 5.2.2, it is inferred among these previous fixed time step simulations performed with $\{\mathcal{M}_x, \mathcal{M}_y\}$ that a prediction reliability R_f exactly equal to prediction reliability goal of $R_{lim} = 0.9$ would have been obtained between Δt = 5s and Δt = 10s that had associated costs $800u$ and $400u$, respectively. Figure 5.8 goes on to demonstrate that maintaining a prediction reliability $R_{lim} = 0.9$ strictly within the computational budget of $400u$ is infeasible for this problem. As such, the L_R – optimized simulation maintains R_{lim} and exceeds the over-restrictive cost limit C_{lim} by 55% to a $C_f = 620u$ and the optimized time step is shown to slowly decrease throughout the analysis from approximately 15s to 5s.

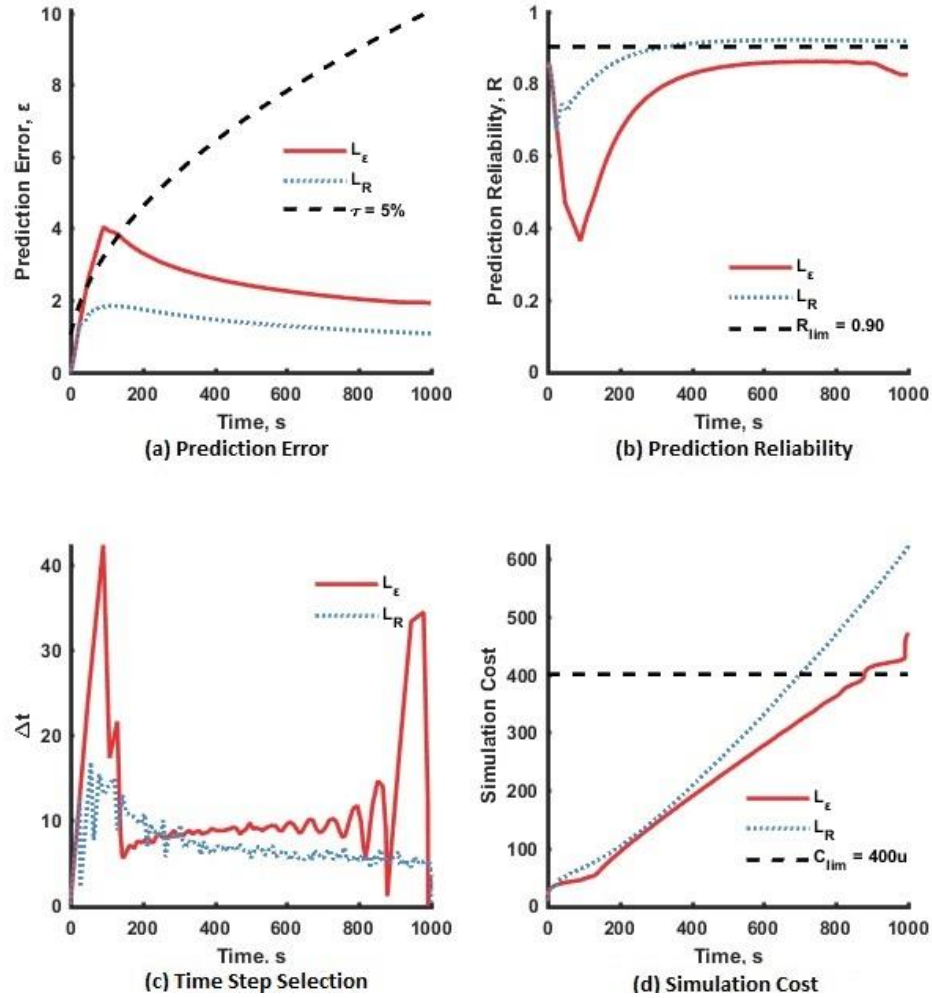

Figure 5.8. Comparison of L_ε and L_R for fixed model $\{M_x, M_y\}$, $[\tau = 5\%, C_{lim} = 400u, R_{lim} = 0.9]$

Conversely to the simulation optimized with L_R, the L_ε - optimization simulation exceeds the cost limit C_{lim} by 18% yet does not reach the desired prediction reliability goal. After comparing the performances of both L_ε and L_R-optimized simulations, the reliability-based loss function L_R is used going forward.

5.3.4 Case 3: Combined Coupling and Temporal Fidelity Selection

The reliability-based loss function L_R was used for the combined optimization to choose between the two coupling scenarios, $\{\mathcal{M}_x,\mathcal{M}_y\}$ and $\{\emptyset_x,\mathcal{M}_y\}$, and well as the time-step for each. Figure 5.9 illustrates how simulation cost vs. prediction reliability tradeoffs impact both coupling

and time-step selection when both C_{lim} and R_{lim} are restricted. Also, the effects of random sampling for uncertainty propagation are shown here, where trends observed for C_{lim}=800u differ from the C_{lim} = 600u and 1000u cases.

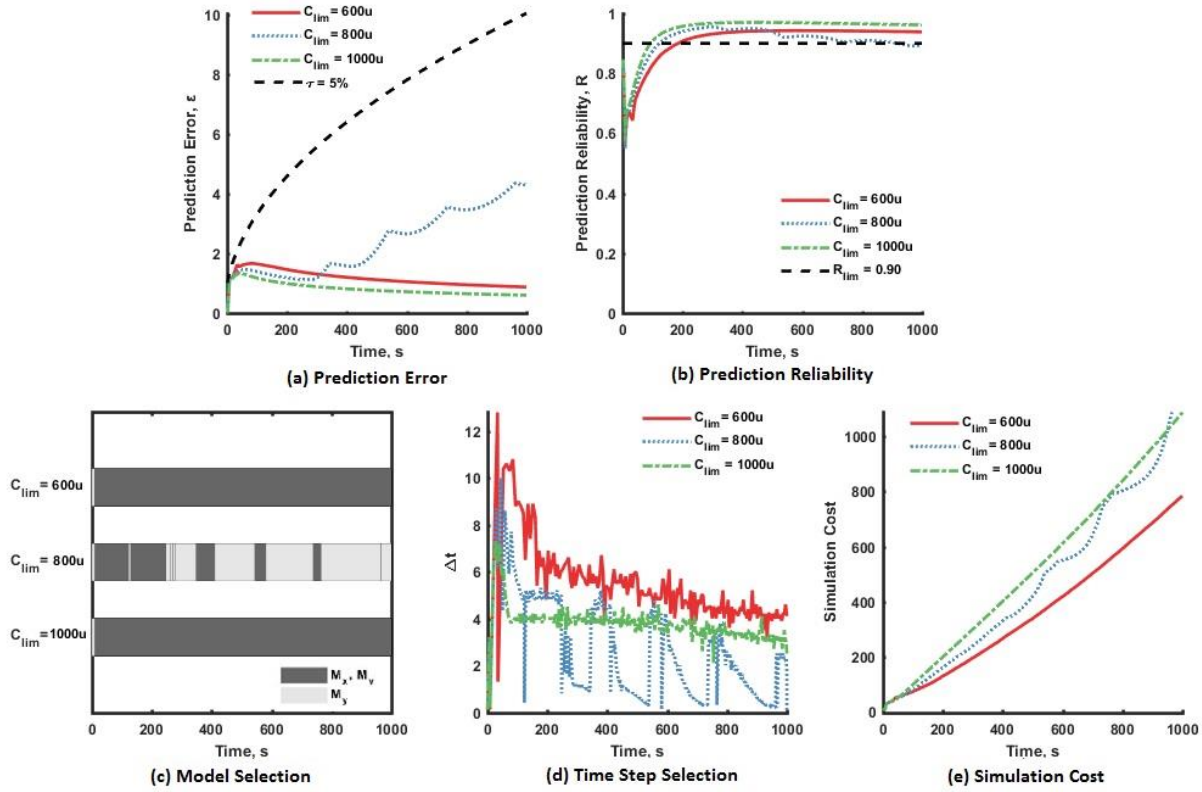

Figure 5.9. L_R for variable coupling and time step
[τ = 5%, C_{lim} = 600, 800, 1000u, and R_{lim} = 0.9]

In Figure 5.9d, for the case when C_{lim}=800u is observed that each model exhibited its own trends in the time step selection, where the two-physics model opted for the larger time steps and the single-physics model trended toward smaller, as expected. In addition, Figure 5.9b shows that this optimization prioritized prediction reliability over the cost objective since C_{lim} was exceeded.

5.4 Application: Aerothermal Coupling and Time-Step Selection

Section 5.4 optimizes a coupled aerothermal simulation through time using the reliability-based loss function formulation for optimization. Recall from Chapters 3 and 4, the aerodynamic heating

errors $\delta_{Q,t}$ (from Eq. (3.32)) from Eckert's reference temperature and 1-D heat transfer discrepancy parameters $\delta_{T,t}$ (from Eq. (3.33)) propagate through to the prediction at each iteration and are functions of the changing model inputs. These model discrepancy parameters were calibrated using fixed coupling $\{Q,T_w\}$ and a fixed time step $\Delta t = 0.05$s. Using these posterior model discrepancies, the reliability-based loss function for model coupling and time-step selection from Section 5.3.1 is applied to the two-discipline aerothermal prediction through a 3 second time history at the leading edge of the Run 30 dome.

Sections 5.4.1 through 5.4.3 illustrate the performance vs cost tradeoffs when optimizing a) the coupling fidelity under a fixed time step, b) the temporal fidelity for a fixed model coupling scenario and c) combined coupling and temporal fidelity.

5.4.1 Case 1: *Coupling Fidelity Selection (Fixed Time Step)*

Table 5.2 lists the costs of each individual model considered in the aerothermal example as well as the combined simulation cost. While presented in terms of the cost units *u*, these do reflect the true cost proportions where Eckert's reference temperature method is three times as expensive as 1-D heat transfer. Figure 5.10 shows the L_R-optimized simulations at a fixed time step $\Delta t = 0.05$s when C_{lim} are imposed as percentages – 25%, 50%, and 100% – of the reference simulation cost C_0. The reference cost C_0 is the cost of using the two-physics prediction$\{Q,T_w\}$ at $\Delta t = 0.05$s over a 3 second time history, which is the same as the cost of the calibrated simulation.

Table 5.2. Aerothermal model costs

	Cost
Eckert's reference temperature method (Q)	3u
1-D Heat Transfer (T_w)	1u
Combined (Q and T_w)	4u
Reference simulation cost, C_0	240u

Figure 5.10 shows the L_R-optimized simulations at a fixed time step Δt = 0.05s when the cost limits imposed as percentages – C_{lim} is 25%, 50%, and 100% of the reference simulation cost C_0. First, the imposed cost limit C_{lim} of 25% of C_0 was met at 60u, however, a sacrifice to prediction reliability was observed since the fully coupled analysis was only evaluated 6 times over the 3 second analysis. Under the computational budgets of 50% and 100% of C_0, however, both cost and prediction reliability targets were met.

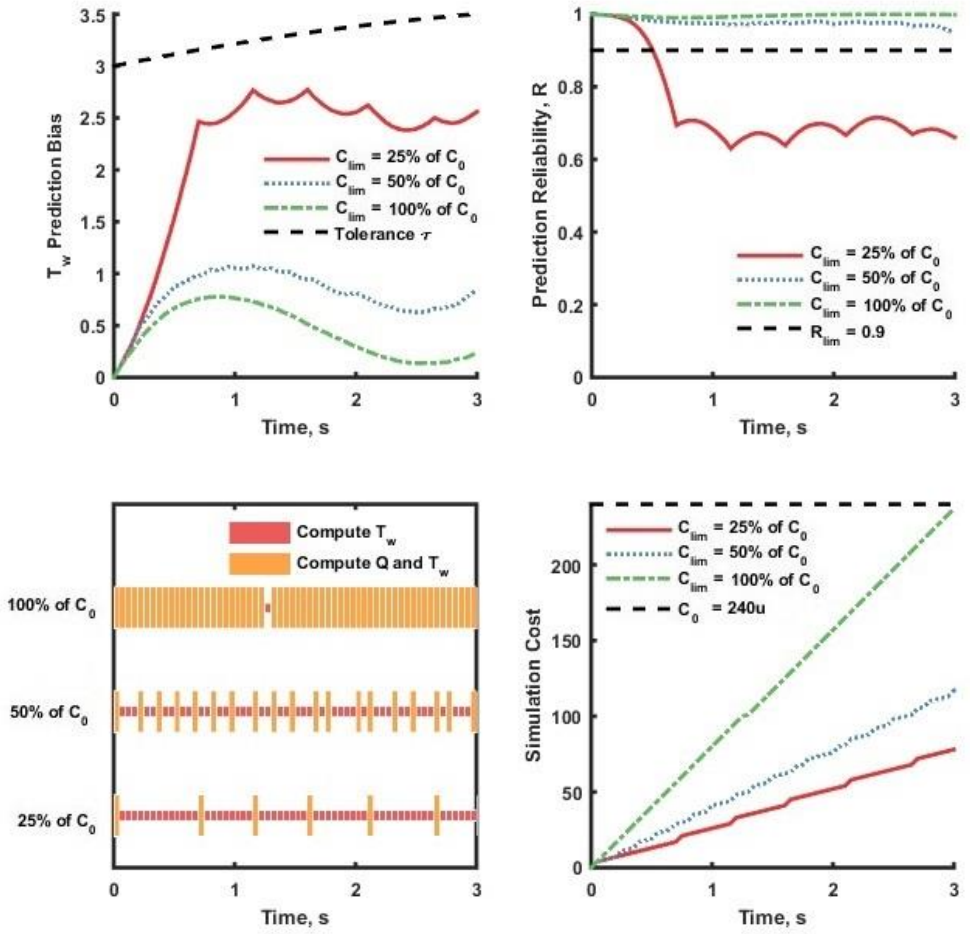

Figure 5.10. Aerothermal optimization for variable coupling and fixed time step [τ = 1%, C_{lim} = 20%, 50%, and 100% of C_0, and R_{lim} = 0.9]

5.4.2 Case 2: Temporal Fidelity Selection (Fixed Coupling)

Figure 5.11 compares the L_R-optimized simulations for a fixed model coupling and a variable time step. First, it was observed that all computational budget criteria were met under each computational budget – 25%, 50%, and 100% of the reference simulation cost C_0 = 240. Intuitively, the more restrictive cost limit led the optimization to select a more coarsely discretized temporal fidelity.

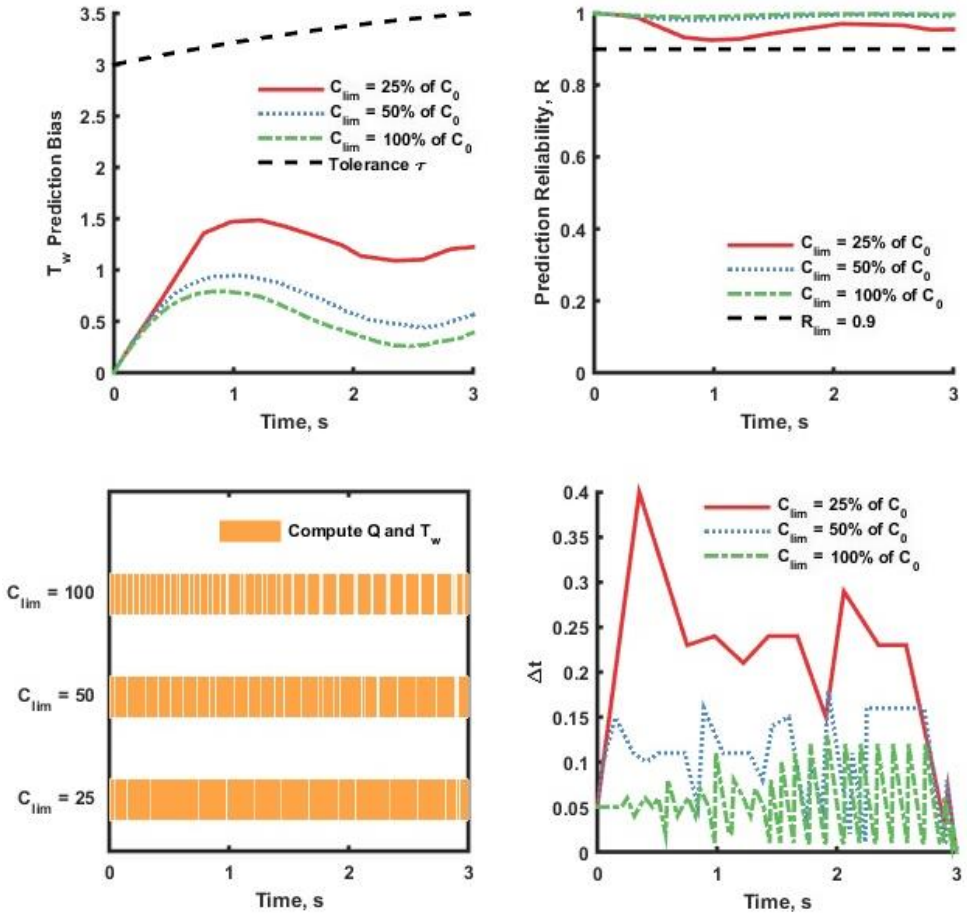

Figure 5.11. Aerothermal optimization for fixed coupling and and variable time step [τ = 5%, C_{lim} = 20%, 50%, and 100% of C_0, and R_{lim} = 0.9]

5.4.3 Case 3: Combined Coupling and Temporal Fidelity Selection

Figure 5.12 demonstrates the developed methodology for the combined coupling and temporal fidelity selection. First, it was observed that all simulations maintained the computational budget, however, however, the sacrifice to model reliability when integrated with a lower-fidelity model is observed Figure 5.11.

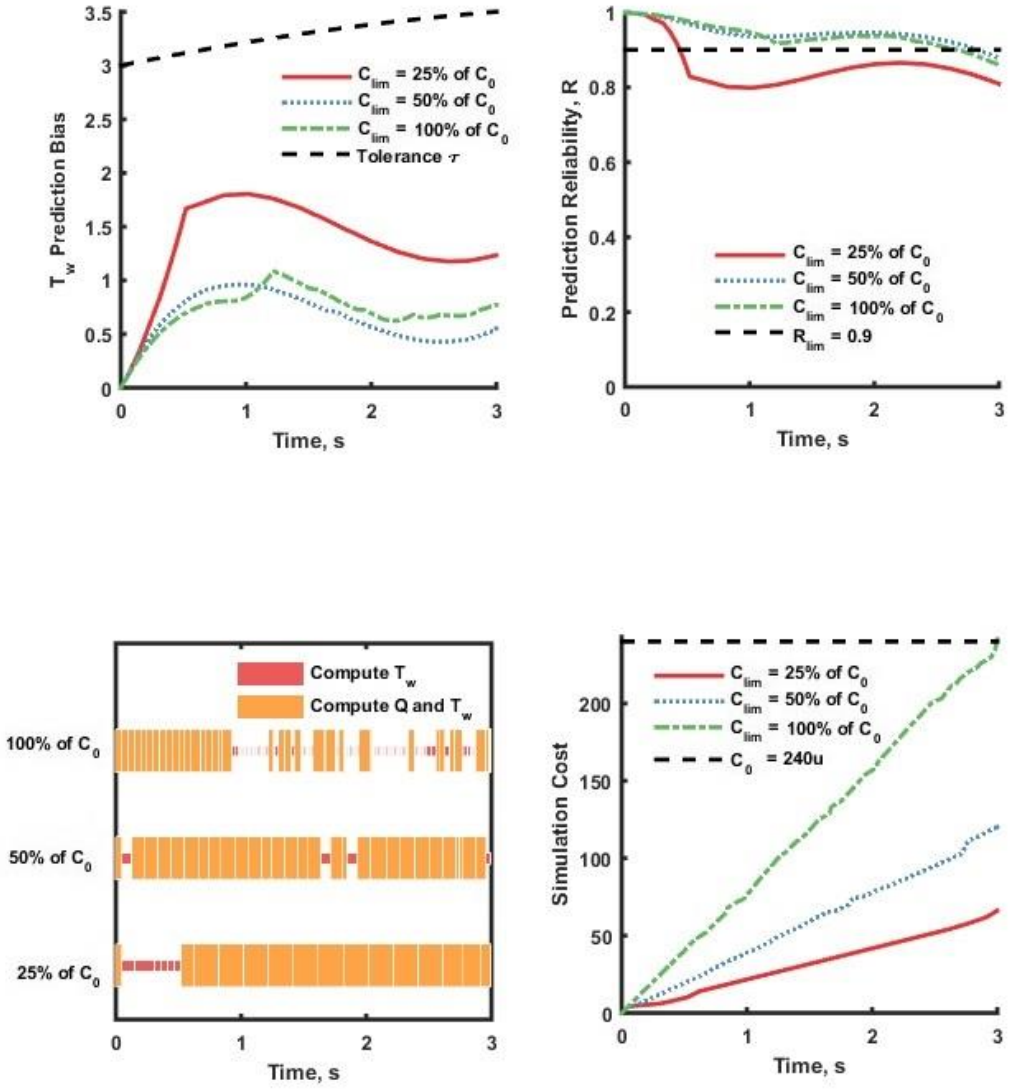

Figure 5.12. Aerothermal optimization for variable coupling and time step [τ = 5%, C_{lim} = 25%, 50%, and 100% of C_0, and R_{lim} = 0.9]

5.5 Conclusion

Two loss functions were formulated – accuracy-based and reliability-based – and compared on their ability to select between coupling fidelity and time steps while maintaining the balance of simulation cost vs. prediction reliability. First, the reliability based loss function exhibit better convergence properties for time-dependent problems. This reliability-based loss function was applied to the aerothermal relationship for coupling and time step selection where optimized predictions were shown to maintain reliability under budget constraints for many tested cases.

Further investigation into the tradeoffs between computational cost and prediction reliability for the fully-coupled aerothermoelastic problem are needed. For instance, the time-scale needed for aeroelastic predictions are much finer than thermal analyses since heat transfer through a structure is a slower-moving phenomenon than structural dynamics. Further advancements could include UQ efforts (i.e., surrogate modeling, calibration, sensitivity analysis) that include structural dynamics analysis which was not addressed in this dissertation.

CHAPTER 6

CONCLUSION

6.1 Summary of Contributions

Motivated by the USAF's need for quantifying and improving prediction confidence in aerothermoelastic simulations, this dissertation addressed several uncertainty quantification challenges regarding the calibration and confidence assessment of multidisciplinary models. The contributions herein addressed both the forward problem of prediction and the inverse problem of model calibration where the challenges are the following: 1) computational expense of multidisciplinary simulations, 2) error accumulation across multiple models and over time, and 3) uncertainty due to the availability of limited data.

Thus, this dissertation expanded the current state of the art in UQ to multidisciplinary analyses, where previous research predominantly focused on single-physics problems. Of particular interest was the challenge of model calibration when errors and uncertainty aggregate across disciplinary models. Therefore, this dissertation addressed model error isolation for the inverse problem of uncertainty reduction and minimizing the computational expense of multidisciplinary model calibration and prediction while maintaining prediction confidence.

In Chapter 2, the computational expense of multidisciplinary model calibration was improved with a segmented Bayesian model calibration strategy. This methodology was developed to guide uncertainty reduction efforts when single- and combined-effect experiments are to be integrated for model calibration. When applied, the segmented calibration strategy was shown to better isolate sources of uncertainty within the calibration segments and the data and model characteristics that resulted in limited sacrifice to downstream prediction confidence were identified. The methodology was used to calibrate aerodynamic pressure and heat flux model discrepancies.

Chapter 3 addressed model discrepancy in coupled, time-dependent simulations where the challenge is isolating sources of uncertainty with data on one output QoI. In Objective 2, three model discrepancy resolutions were developed (global, step-wise, and partitioned) to account for the sources of model error that aggregate through the coupling interfaces and through time. Effective parameterization of the model discrepancy was addressed for the application problem of calibrating aerodynamic heating and 1-D heat transfer models, and the partitioned discrepancy approach was shown to better capture data nonlinearities and demonstrated increased prediction reliability in extrapolation.

The computational challenges of identifying significant sources of uncertainty using global sensitivity analysis (GSA) were addressed in Chapter 4. For independent variables, GSA was made more efficient using Sobol' sequences for sampling the input parameter space with an importance sampling-based kernel regression method (ISK). Sensitivities were shown to asymptotically converge and resulted in fewer model evaluations for independent variables using Sobol' sequences compared to pseudo-random number generators. Efficient GSA methods were then further extended in this chapter to dependent variables sensitivity analysis. Specifically, the methodology aided post-calibration GSA by using existing input-output samples from model calibration directly and did not require additional model evaluations.

Finally, Chapter 5 addressed model selection in multidisciplinary analyses to identify the necessary coupling and temporal fidelities needed to maintain prediction confidence under a restricted computational budget. Two loss functions were formulated – accuracy-based and reliability-based, where the reliability-based loss function demonstrated better convergence properties for time-dependent problems. The developed reliability-based loss function for model selection was applied to the aerothermal relationship for coupling and time step selection where

the optimized predictions were shown to maintain reliability under budget constraints for many tested cases.

6.2 Future Work

This dissertation was limited to the aerothermal model components (aerodynamic pressure, aerodynamic heating, and heat transfer in Figure 1.1) due to the limitation of the Glass and Hunt experiments used for calibration (only pressure and heat flux were measured for a rigid dome, no structural deformation). Further research that extends the methods developed herein to the fully-coupled aerothermoelastic analysis requires data on structural deformation and dynamics.

Further, other higher fidelity, reduced order, and surrogate models for each discipline have been developed in-house by the Structural Sciences Center at AFRL and can be integrated into aerothermoelastic analysis. These alternative modeling choices can be integrated into a multi-fidelity framework to further advance the model selection objective. The combination of multi-fidelity modeling and additional aerothermoelastic data will significantly advance the methodologies proposed in this dissertation towards application to hypersonic vehicle components.

APPENDIX A

ANALYTICAL DERIVATIONS OF SEGMENTED AND SIMULTANEOUS CALIBRATION POSTERIORS

The contents of this Appendix are the additional derivations that were used to produce the results in Section 2.6.

Analytical expressions for the posterior distributions for segmented and simultaneous calibration of θ_1 and θ_2 are derived below using the assumption of conjugate distributions [55]. Both the priors and posteriors of the calibration parameters are assumed to be normal only for this analytical example.

A.1. Segmented Calibration

The segmented Bayesian calibration of θ_1 is formulated in Eq. (A.1) and the likelihood $L(\theta_1)$ in Eq. (A.2) is derived by assuming normally distributed measurement errors with zero mean and known variance $V(y_{D1})$[88]. Note that the prediction from the first model is $y_1(\theta_1) = \theta_1$, as in Section 2.6.

$$\pi(\theta_1 \mid y_{D1}) \propto L(\theta_1)\pi(\theta_1) \tag{A.1}$$

$$L(\theta_1) \propto \prod_{i=1}^{N_1} \pi(y_{D1,i} \mid \theta_1) \propto \exp\left[-\frac{1}{2} \frac{(\theta_1 - \overline{y_{D1}})^2}{\sigma_{D1}^2 / N_1} \right] \tag{A.2}$$

Multiplying Eq. (A.2) by the normal prior distribution of θ_2 with mean $E(\theta_2)$ and variance $V(\theta_2)$, Eqs. (A.3) and (A.4) are the posterior statistics of θ_1 with mean $E[\theta_1|y_{D1}]$ and posterior variance $Var[\theta_1|y_{D1}]$. The posterior statistics are derived by assuming normal conjugate prior and posterior distributions.

$$E[\theta_1 \mid y_{D1}] = \frac{N_1 E(y_{D1}) V(\theta_1) + E(\theta_1) V(y_{D1})}{N_1 V(\theta_1) + V(y_{D1})} \tag{A.3}$$

$$V[\theta_1 \mid y_{D1}] = \frac{V(\theta_1) V(y_{D1})}{V(\theta_1) + V(y_{D1})} \tag{A.4}$$

The prediction using the next model is $y_2(\theta_1, \theta_2) = c\theta_1 + \theta_2$, and the distribution of θ_1 is now aleatory and not updated from Eqs. (A.3) and (A.4) with the new data y_2. The Bayesian formulation and likelihood function for the subsequent calibration of θ_2 with y_{D2} are shown in Eq. (A.5) and (A.6), respectively. and thus the likelihood $L(\theta_2)$ in Eq. (A.6) includes the posterior distribution of $\theta_1|D_1$ with statistics from Eqs. (A.3) and (A.4).

$$\pi(\theta_2 \mid y_{D2}) \propto L(\theta_2)\pi(\theta_2) \tag{A.5}$$

$$\begin{aligned} L(\theta_2) &\propto \prod_{i=1}^{N_2} \pi(y_{D2,i} \mid c\theta_{1|D1} + \theta_2) \\ &\propto \exp\left[-\frac{N_2}{2} \frac{(cE[\theta_1 \mid y_{D1}] + \theta_2 - E(y_{D2}))^2}{c^2 V[\theta_1 \mid y_{D1}] + V(y_{D2})} \right] \end{aligned} \tag{A.6}$$

Multiplying Eq. (A.6) by the normal prior of θ_2, the posterior mean $E[\theta_2|y_{D2}]$ and posterior variance $Var[\theta_2|y_{D2}]$ are derived as in Eqs. (A.7) and (A.8), respectively, similar to the derivations of Eqs. (A.3) and (A.4).

$$E[\theta_2 \mid y_{D2}] = \frac{N_2 V(\theta_2)\left[E(y_{D2}) - cE[\theta_1 \mid y_{D1}]\right] + E(\theta_2)\left[c^2 V[\theta_1 \mid y_{D1}] + V(y_{D2})\right]}{N_2 V(\theta_2) + c^2 V[\theta_1 \mid y_{D1}] + V(y_{D2})} \tag{A.7}$$

$$V[\theta_2 \mid y_{D2}] = \frac{N_2 V(\theta_2)\left[c^2 V[\theta_1 \mid y_{D1}] + V(y_{D2})\right]}{N_2 V(\theta_2) + c^2 V[\theta_1 \mid y_{D1}] + V(y_{D2})} \tag{A.8}$$

The posterior statistics in Eqs. (A.7) and (A.8) are derived assuming there is zero correlation between parameters θ_1 and θ_2, which is an essential feature of segmented calibration. Next, the posterior statistics of θ_1 and θ_2 are propagated to the prediction of y_2 as in Eq. (A.9) and (A.10).

$$E[y_2]_{seg} = cE[\theta_1 \mid y_{D1}] + E[\theta_2 \mid y_{D1}] \tag{A.9}$$

$$V[y_2]_{seg} = c^2 V[\theta_1 \mid y_{D1}] + V[\theta_2 \mid y_{D2}] \tag{A.10}$$

A.2. Simultaneous Calibration

The simultaneous formulation of Bayesian calibration of parameters θ_1 and θ_2 with data y_{D1} and y_{D2} is shown in Eq. (A.11) and the joint likelihood is formulated in Eq. (A.12).

$$\pi(\theta_1, \theta_2 \mid y_{D1}, y_{D2}) \propto L(\theta_1, \theta_2)\pi(\theta_1, \theta_2) \tag{A.11}$$

$$\begin{aligned} L(\theta_1, \theta_2) &\propto \pi(y_{D1}, y_{D2} \mid \theta_1, \theta_2) \\ &\propto \pi(y_{D1} \mid \theta_1)\pi(y_{D2} \mid c\theta_1 + \theta_2) \\ &\propto \exp\left[-\frac{N_1}{2} \frac{(\theta_1 - E(y_{D1}))^2}{V(y_{D1})} \right] \exp\left[-\frac{N_2}{2} \frac{(c\theta_1 + \theta_2 - E(y_{D2}))^2}{V(y_{D2})} \right] \end{aligned} \tag{A.12}$$

Using the result in Eq. (A.2), the likelihood in Eq. (A.12) is multiplied by the normal prior of θ_1 and θ_2. The joint posterior means and variances of θ_1 and θ_2 and correlation coefficient between θ_1 and θ_2 can be derived as follows.

$$\begin{aligned} &E[\theta_1 \mid y_{D1}, y_{D2}] = \\ &\quad \frac{N_2 c V(\theta_1) V(y_{D1}) \left[E(y_{D2}) - E(\theta_2) \right]}{N_2 c^2 V(\theta_1) V(y_{D1}) + \left[N_2 V(\theta_2) + V(y_{D2}) \right] \left[N_1 V(\theta_1) + V(y_{D1}) \right]} + \\ &\quad \frac{\left[N_2 V(\theta_2) + V(y_{D2}) \right] \left[N_1 E(y_{D2}) V(\theta_1) + E(\theta_1) V(y_{D1}) \right]}{N_2 c^2 V(\theta_1) V(y_{D1}) + \left[N_2 V(\theta_2) + V(y_{D2}) \right] \left[N_1 V(\theta_1) + V(y_{D1}) \right]} \end{aligned} \tag{A.13}$$

$$V[\theta_1 \mid y_{D1}, y_{D2}] = \frac{V(\theta_1) V(y_{D1}) \left[N_2 V(\theta_2) + V(y_{D2}) \right]}{N_2 c^2 V(\theta_1) V(y_{D1}) + \left[N_2 V(\theta_2) + V(y_{D2}) \right] \left[N_1 V(\theta_1) + V(y_{D1}) \right]} \tag{A.14}$$

$$\begin{aligned} &E[\theta_2 \mid y_{D1}, y_{D2}] = \\ &\quad \frac{N_2 V(\theta_2) N_1 V(\theta_1) (E(y_{D1}) - cE(y_{D2})) + N_2 V(\theta_1) V(y_{D1}) (E(y_{D2}) - cE(\theta_1))}{N_2 c^2 V(\theta_1) V(y_{D1}) + \left[N_2 V(\theta_2) + V(y_{D2}) \right] \left[N_1 V(\theta_1) + V(y_{D1}) \right]} + \\ &\quad \frac{E(\theta_2) \left[N_2 c^2 V(\theta_1) V(y_{D1}) + V(y_{D2}) (N_1 V(\theta_1) + V(y_{D1}) \right]}{N_2 c^2 V(\theta_1) V(y_{D1}) + \left[N_2 V(\theta_2) + V(y_{D2}) \right] \left[N_1 V(\theta_1) + V(y_{D1}) \right]} \end{aligned} \tag{A.15}$$

$$V[\theta_2 \mid y_{D1}, y_{D2}] = \frac{N_2 V(\theta_2)V(\theta_1)V(y_{D1}) + V(y_{D2})V(\theta_2)(N_1 V(\theta_1) + V(y_{D1}))}{N_2 c^2 V(\theta_1)V(y_{D1}) + \left[N_2 V(\theta_2) + V(y_{D2})\right]\left[N_1 V(\theta_1) + V(y_{D1})\right]} \quad \text{(A.16)}$$

$$\rho(\theta_1, \theta_2 \mid y_{D1}, y_{D2}) = \frac{-N_2 c\sqrt{V(\theta_2)V(\theta_1)V(y_{D1})}}{\sqrt{\left[N_2 c^2 V(\theta_1)V(y_{D1}) + (N_1 V(\theta_1) + V(y_{D1}))\right]\left[N_2 V(\theta_2) + V(y_{D2})\right]}} \quad \text{(A.17)}$$

Next, the posterior statistics of θ_1 and θ_2 are propagated to the prediction of y_2 to obtain

$$E[y_2]_{sim} = cE[\theta_1 \mid y_{D1}, y_{D2}] + E[\theta_2 \mid y_{D1}, y_{D2}] \quad \text{(A.18)}$$

$$V[y_2]_{sim} = c^2 V(\theta_1 \mid y_{D1}, y_{D2}) + V(\theta_2 \mid y_{D1}, y_{D2}) + 2c\rho(\theta_1, \theta_2 \mid y_{D1}, y_{D2})\sqrt{V[\theta_1 \mid y_{D1}, y_{D2}]V[\theta_2 \mid y_{D1}, y_{D2}]} \quad \text{(A.19)}$$

A.3. Segmented Calibration with Shared Parameters θ_1

In this case, θ_1 is a shared parameter whose posterior from the first calibration is the prior for the second calibration. The Bayesian formulation and the likelihood are shown in Eqs. (A.20) and (A.21).

$$\pi(\theta_1, \theta_2 \mid y_{D1}, y_{D2}) \propto L(\theta_1, \theta_2)\pi(\theta_2)\pi(\theta_1 \mid y_{D1}) \quad \text{(A.20)}$$

$$L(\theta_1, \theta_2) \propto \exp\left[-\frac{1}{2}\frac{N_2(E(y_{D1} - c\theta_1 - \theta_2)^2}{V(y_{D2})}\right] \quad \text{(A.21)}$$

Multiplying Eq. (A.21) by the normal prior of θ_2 and $\pi(\theta_1|y_{D1})$, the joint posterior means and variances of θ_1 and θ_2 and correlation coefficient between θ_1 and θ_2 are derived as follows.

$$E[\theta_1 \mid y_{D1}, y_{D2}] = \frac{N_2 cV[\theta_1 \mid y_{D1}]\left[E(y_{D2}) - cE(\theta_2)\right] + E[\theta_1 \mid y_{D1}]\left[N_2 V(\theta_2) + V(y_{D2})\right]}{N_2\left[c^2 V[\theta_1 \mid y_{D1}] + V(\theta_2)\right] + V(y_{D2})} \quad \text{(A.22)}$$

$$V[\theta_1 \mid y_{D1}, y_{D1}] = \frac{V[\theta_1 \mid y_{D1}]\left[N_2 V(\theta_2) + V(y_{D2})\right]}{N_2\left[c^2 V(\theta_1 \mid y_{D2}) + V(\theta_2)\right] + V(y_{D2})} \quad \text{(A.23)}$$

$$E[\theta_2 \mid y_{D1}, y_{D2}] = \frac{V(\theta_2)\left[N_2 c^2 V[\theta_1 \mid y_{D1}] + V(y_{D2})\right]}{N_2\left[c^2 V[\theta_1 \mid y_{D1}] + V(\theta_2)\right] + V(y_{D2})} \tag{A.24}$$

$$V[\theta_2 \mid y_{D1}, y_{D2}] = \frac{V(\theta_2)\left[N_2 V(\theta_2) + V(y_{D2})\right]}{N_2\left[c^2 V[\theta_1 \mid y_{D1}] + V(\theta_2)\right] + V(y_{D2})} \tag{A.25}$$

$$\rho(\theta_1, \theta_2 \mid y_{D1}, y_{D2}) = \frac{-N_2 c\sqrt{V(\theta_2)V(\theta_1 \mid y_{D1})}}{\sqrt{\left[N_2 c^2 V(\theta_1 \mid y_{D1}) + V(\theta_2)\right]\left[N_2 V(\theta_2) + V(y_{D2})\right]}} \tag{A.26}$$

REFERENCES

[1] Rose, L. J., "Air Force Research Laboratory's Focused Long Term Challenges," Dayton, OH, U.S. Air Force Research Laboratory, Munitions Directorate, 2008.

[2] Chona, R., "Assuring aero-structural reliability in an uncertain world (Keynote)," *16th AIAA Non-Deterministic Approaches Conference at AIAA SciTech*, AIAA 2014-0123, 2014.

[3] Bodony, D. J., Zagaris, G., Reichert, A., and Zhang, Q., "Aeroacoustic predictions in complex geometries," *Procedia IUTAM*, Vol. 1, 2010, pp. 234–243. doi: 10.1016/j.piutam.2010.10.025

[4] Culler, A. J., and McNamara, J. J., "Impact of Fluid-Thermal-Structural Coupling on Response Prediction of Hypersonic Skin Panels," *AIAA Journal*, Vol. 49, No. 11, 2011, pp. 2393–2406. doi: 10.2514/1.J050617

[5] Tuegel, E. J., Ingraffea, A. R., Eason, T. G., and Spottswood, S. M., "Reengineering aircraft structural life prediction using a digital twin," *International Journal of Aerospace Engineering*, Vol. 2011, 2011. doi: 10.1155/2011/154798

[6] "NASA Technology Roadmaps - TA 11: Modeling, Simulation, Information Technology and Processing Roadmap," 2015.

[7] Ostoich, C. M., "Aerothermal and aeroelastic response prediction of aerospace structures in high-speed flows using direct numerical simulation," University of Illinois, Urbana-Champaign, 2013.

[8] Crowell, A. R., Mcnamara, J. J., and Miller, B. a, "Surrogate Based Reduced-Order Aerothermodynamic Modeling for Structural Response Prediction at High Mach Numbers," *52nd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics & Materials Conference*, AIAA 2011-2014, Apr. 2011, No. April 2011.

[9] Chen, X., Liu, L., Long, T., and Yue, Z., "A reduced order aerothermodynamic modeling framework for hypersonic vehicles based on surrogate and POD," *Chinese Journal of Aeronautics*, Vol. 28, No. 5, 2015, pp. 1328–1342. doi: 10.1016/j.cja.2015.06.024

[10] Beléndez, T., Pérez-Polo, M., Neipp, C., and Beléndez, A., "Numerical and Experimental Analysis of Large Deflections of Cantilever Beams Under a Combined Load," *Physica Scripta*, Vol. T118, No., 2005, pp. 61–65. doi: 10.1238/Physica.Topical.118a00061

[11] Brouwer, K., Crowell, A. R., and McNamara, J. J., "Rapid Prediction of Unsteady Aeroelastic Loads in Shock-Dominated Flows," *56th AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, AIAA 2015-0687, 2015.

[12] Glaessgen, E. H., and Stargel, D. S., "The Digital Twin Paradigm for Future NASA and U.S. Air Force Vehicles," *53rd Structures, Structural Dynamics and Materials Conference - Special Session: Digital Twin*, No. April, 2012, pp. 1–14. doi: 10.2514/6.2012-1818

[13] Zuchowski, B., "Air Vehicle Integration and Technology Research (AVIATR) Task Order 0015 : Predictive Capability for Hypersonic Structural Response and Life Prediction , Phase 1 - Identification of Knowledge Gaps," AFRL-RB-WP-TR-2010-3069, 2010.

[14] Tzong, G., Jacobs, R., and Liguore, S., "Air Vehicle Integration and Technology Research (AVIATR) Task Order 0015 : Predictive Capability for Hypersonic Structural Response and Life Prediction : Phase 1-Identification of Knowledge Gaps, Volume 1 - Nonproprietary Version," AFRL-RB-WP-TR-2010-3068,V1, 2010.

[15] Liguore, S. L., Pitt, D. M., Thomas, M. J., and Gurtowski, N., "Air Vehicle Integration and Technology Research (AVIATR) Delivery Order 0013: Nonlinear, Low-Order/Reduced-Order Modeling Applications and Demonstration Salvatore," AFRL-RB-WP-TR-2011-3102, 2011.

[16] Spottswood, S. M., Eason, T. G., and Beberniss, T., "Influence of shock-boundary layer interactions on the dynamic response of a flexible panel," *International Conference on Noise and Vibration Engineering (ISMA)*, 2012, pp. 603–616.

[17] Spottswood, S. M., Beberniss, T. J., and Eason, T. G., "Full-field, dynamic pressure and displacement measurements of a panel excited by shock boundary-layer interaction," AFRL-RQ-WP-TP-2015-0046, 2015.

[18] Deveikis, W. D., and Hunt, L. R., "Loading and Heating of a Large Flat Plate at Mach 7 in the Langley 8-Foot High-Temperature Structures Tunnel," NASA TN D-7275, 1973.

[19] Deveikis, W. D., Bruce, W. E., and Karns, J. R., "Techniques for Aerothermal Tests of Large, Flightweight, Thermal Protection Panels in a Mach 7 Wind Tunnel," NASA TM X-71983, 1974.

[20] Glass, C. E., and Hunt, L. R., "Aerothermal tests of Spherical Dome Protuberances on a Flat Plate at a Mach number of 6.5," NASA TP-2631, 1986.

[21] Glass, C. E., and Hunt, L. R., "Aerothermal Tests of Quilted Dome Models on a Flat Plate at a Mach Number of 6.5," NASA TP-2804, 1988.

[22] Ostoich, C. M., Bodony, D. J., and Geubelle, P. H., "Fluid-Thermal Response of Spherical Dome Under a Mach 6.59 Laminar Boundary Layer," *AIAA Journal*, Vol. 50, No. 12, 2012, pp. 2791–2808. doi: 10.2514/1.J051634

[23] Culler, A. J., and McNamara, J. J., "Studies on fluid-thermal structural coupling for aerothermoelasticity in hypersonic flow," *AIAA Journal*, Vol. 48, No. 8, 2010, pp. 1721–1738. doi: 10.2514/1.J050193

[24] Blades, E., Shah, P. N., Nucci, M., and Miskovish, S., "Demonstration of Multiphysics Analysis Tools on Representative Hypersonic Vehicle Structures," *54th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, American Institute of Aeronautics and Astronautics, 2013

[25] Smarslok, B. P., and Villanueva, D., “Design of Experiments for Model Calibration of Multi-Physics Systems with Targeted Events of Interest,” AFRL-RQ-WP-TP-2017-0034, 2017.

[26] Sankararaman, S., “Uncertainty Quantification and Integration,” Vanderbilt University, Nashville, TN, 2012.

[27] Anderson Jr., J. D., *Hypersonic and High-Temperature Gas Dynamics, Second Edition*. American Institute of Aeronautics and Astronautics, 2006.

[28] Culler, A. J., and Mcnamara, J. J., “Fluid-thermal-structural modeling and analysis of hypersonic structures under combined loading,” *52nd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials and Co-located Conferences*, AIAA 2011-1965, Apr. 2011.

[29] Blevins, R. D., Bofilios, D., Holehouse, I., Hwa, V. W., Tratt, M. D., Laganelli, A. L., Pozefsky, P., and Pierucci, M., “Thermo-Vibro-Acoustic Loads and Fatigue of Hypersonic Flight Vehicle Structure,” AFRL-RB-WP-TR-2009-3139, 2009.

[30] Kontinos, D., “Coupled Thermal Analysis Method with Application to Metallic Thermal Protection Panels,” *Journal of Thermophysics and Heat Transfer*, Vol. 11, No. 2, Apr. 1997, pp. 173–181. doi: 10.2514/2.6249

[31] Falkiewicz, N., “Enhanced Modal Solutions for Structural Dynamics in Aerothermoelastic Analysis,” *Aerospace*, AIAA 2011-1963, Apr. 2011.

[32] Holt, A., and Zartarian, G., “Piston Theory: A New Aerodynamic Tool for the Aeroelastician,” *Journal of the Aeronautical Sciences*, Vol. 23, No. 12, 1956, pp. 1109–1118. doi: 10.2514/8.3740

[33] Eckert, E. R. G., “Engineering Relations for Heat Transfer and Friction in High-Velocity Laminar and Turbulent Boundary-Layer Flow over Surfaces with Contstant Pressure and Temperature,” *Transactions of the ASME*, Vol. 78, No. 6, 1956, pp. 1273–1283.

[34] McNamara, J. J., “Aeroelastic and Aerothermoelastic Analysis in Hypersonic Flow : Past , Present , and Future,” *AIAA Journal*, Vol. 49, No. 6, 2013, pp. 23–26. doi: 10.2514/1.J050882

[35] Crowell, A. R., Miller, B. a, and Mcnamara, J. J., “Robust and Efficient Treatment of Temperature Feedback in Fluid-Thermal-Structural Analysis,” *AIAA Journal*, Vol. 52, No. 11, 2014, pp. 1–48. doi: 10.2514/6.2013-1663

[36] Liguore, S. L., and Tzong, G., “Identification of Knowledge Gaps in the Predictive Capability for Response and Life Prediction of Hypersonic Vehicle Structures,” *52nd AIAA/ASME/AHS/ASC Structures, Structural Dynamics and Materials Conference*, AIAA Paper 2011-1961, Apr. 2011, pp. 1–9.

[37] Hu, Z., and Du, X., "Time-dependent reliability analysis with joint upcrossing rates," *Structural and Multidisciplinary Optimization*, Vol. 48, No. 5, 2013, pp. 893–907. doi: 10.1007/s00158-013-0937-2

[38] Sparkman, D., Garza, J., Millwater, H., and Smarslok, B., "Importance Sampling-based Post-processing Method for Global Sensitivity Analysis," *15th*.

[39] Der Kiureghian, A., "Analysis of structural reliability under parameter uncertainties," *Probabilistic Engineering Mechanics*, Vol. 23, No. 4, Oct. 2008, pp. 351–358. doi: 10.1016/j.probengmech.2007.10.011

[40] Haldar, A., and Mahadevan, S., *Probability, Reliability and Statistical Methods in Engineering Design*. Wiley: New York, 2000.

[41] Urbina, A., Mahadevan, S., and Paez, T. L., "Quantification of margins and uncertainties of complex systems in the presence of aleatoric and epistemic uncertainty," *Reliability Engineering and System Safety*, Vol. 96, No. 9, 2011, pp. 1114–1125. doi: 10.1016/j.ress.2010.08.010

[42] Mahadevan, S., Zhang, R., and Smith, N., "Bayesian networks for system reliability reassessment," *Structural Safety*, Vol. 23, No. 3, 2001, pp. 231–251. doi: 10.1016/S0167-4730(01)00017-0

[43] Wang, S., Chen, W., and Tsui, K., "Bayesian Validation of Computer Models," *Technometrics*, Vol. 51, No. 4, 2009, pp. 439–451. doi: 10.1198/TECH.2009.07011

[44] Roy, C. J., and Oberkampf, W. L., "A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing," *Computer Methods in Applied Mechanics and Engineering*, Vol. 200, No. 25–28, 2011, pp. 2131–2144. doi: 10.1016/j.cma.2011.03.016

[45] Sankararaman, S., Ling, Y., and Mahadevan, S., "Uncertainty quantification and model validation of fatigue crack growth prediction," *Engineering Fracture Mechanics*, Vol. 78, No. 7, 2011, pp. 1487–1504. doi: 10.1016/j.engfracmech.2011.02.017

[46] Li, C., and Mahadevan, S., "Role of calibration, validation, and relevance in multi-level uncertainty integration," *Reliability Engineering and System Safety*, Vol. 148, Apr. 2016, pp. 32–43. doi: 10.1016/j.ress.2015.11.013

[47] Sankararaman, S., and Mahadevan, S., "Integration of Model Verification, Validation, and Calibration for Uncertainty Quantification in Engineering Systems," *Reliability Engineering and System Safety*, Vol. 138, Jun. 2015, pp. 194–209. doi: 10.1016/j.ress.2015.01.023

[48] Mullins, J., Li, C., Mahadevan, S., and Urbina, A., "Optimal Selection of Calibration and Validation Test Samples Under Uncertainty," *Model Validation and Uncertainty Quantification, Volume 3: Proceedings of the 32nd IMAC, A Conference and Exposition on*

Structural Dynamics, H. S. Atamturktur, B. Moaveni, C. Papadimitriou, and T. Schoenherr, Eds. Springer International Publishing: Cham, 2014, pp. 391–401

[49] Jeffreys, H., "An invariant form for the prior probability in estimation problems.," *Royal Society of London. Series A: Mathematical and physical sciences*, Vol. 186, No. 1007, Sep. 1946, pp. 453–461. doi: 10.1098/rspa.1946.0056

[50] Kullback, S., and Leibler, R. A., "On Information and Sufficiency," *The Annals of Mathematical Statistics*, Vol. 22, No. 1, 1951, pp. 79–86. doi: 10.1214/aoms/1177729694

[51] Novak, E., and Woźniakowski, H., "Approximation of infinitely differentiable multivariate functions is intractable," *Journal of Complexity*, Vol. 25, No. 4, 2009, pp. 398–404. doi: 10.1016/j.jco.2008.11.002

[52] Hastings, W. K., "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika Vol*, Vol. 57, No. 1, 1970, pp. 97–109.

[53] Casella, G., and George, E. I., "Explaining the Gibbs sampler," *The American Statistician*, Vol. 46, No. 3, Aug. 1992, pp. 167–174. doi: 10.2307/2685208

[54] Neal, R. M., "Slice sampling," *The Annals of Statistics*, Vol. 31, No. 3, 2003, pp. 705–767. doi: 10.1214/aos/1056562461

[55] Ang, A. H.-S., and Tang, W. H., *Probability Concepts in Engineering Planning and Design: Basic principles*. Wiley: New York, 1975.

[56] Ames Research Staff, "Equations, Tables, and Charts for Compressible Flow," NACA Report 1135, 1947.

[57] DeCarlo, E. C., Mahadevan, S., and Smarslok, B. P., "Bayesian Calibration of Aerothermal Models for Hypersonic Air Vehicles," *54th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials and Co-located Conferences*, AIAA 2013-1683, Apr. 2013.

[58] Smarslok, B. P., Culler, A. J., and Mahadevan, S., "Error Quantification and Confidence Assessment of Aerothermal Model Predictions for Hypersonic Aircraft," *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials and Co-located Conferences*, AIAA 2012-1817, Apr. 2012.

[59] De Finetti, B., *Theory of Probability*, 3rd ed., Vol. 1–2. Oxford University Press: New York, NY, 1990.

[60] Oberkampf, W. L., "Practical and Technical Challenges in Verification and Validation," 2012.

[61] Ling, Y., Mullins, J., and Mahadevan, S., "Calibration of multi-physics computational models using Bayesian networks," *Journal of Computational Physics*, 2012, pp. 1–38.

[62] Kennedy, M. C., and O'Hagan, A., "Bayesian calibration of computer models," *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 63, No. 3, 2001, pp. 425–464. doi: 10.1111/1467-9868.00294

[63] Eckert, E. R. G., "Engineering Relations for Heat Transfer and Friction in High Velocity Laminar and Turbulent Boundary Layer Flow over Surfaces with Constant Pressure and Temperature.," *Transactions of the ASME*, Vol. 78, No. 6, 1956, pp. 1273–1283.

[64] DeCarlo, E. C., Mahadevan, S., and Smarslok, B. P., "Bayesian Calibration of Coupled Aerothermal Models Using Time-Dependent Data," *16th AIAA Non-Deterministic Approaches Conference at AIAA SciTech*, AIAA 2014-0123, Jan. 2014.

[65] Rebba, R., and Mahadevan, S., "Model Predictive Capability Assessment Under Uncertainty," *AIAA Journal*, Vol. 44, No. 10, 2006, pp. 2376–2384. doi: 10.2514/1.19103

[66] Rebba, R., and Mahadevan, S., "Computational methods for model reliability assessment," *Reliability Engineering and System Safety*, Vol. 93, Aug. 2008, pp. 1197–1207. doi: 10.1016/j.ress.2007.08.001

[67] Hu, Z., and Mahadevan, S., "Time-Dependent System Reliability Analysis Using Random Field Discretization," *Journal of Mechanical Design*, Vol. 137, No. 10, 2015, p. 101404. doi: 10.1115/1.4031337

[68] Liang, C., and Mahadevan, S., "Bayesian Sensitivity Analysis and Uncertainty Integration for Robust Optimization," *Journal of Aerospace Information Systems*, Vol. 12, No. 1, Nov. 2014, pp. 189–203. doi: 10.2514/1.I010268

[69] Liang, C., and Mahadevan, S., "Multidisciplinary Optimization under Uncertainty Using Bayesian Network," *SAE Int. J. Mater. Manf.*, Vol. 9, No. 2, 2016, pp. 419–429. doi: 10.4271/2016-01-0304

[70] Sobol, I. M., "Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates," *Mathematics and Computers in Simulation*, Vol. 55, No. 1–3, 2001, pp. 271–280. doi: 10.1016/S0378-4754(00)00270-6

[71] Mahadevan, S., and Liang, B., "Error and Uncertainty Quantification and Sensitivity Analysis in Mechanics Computational Models," *International Journal for Uncertainty Quantification*, Vol. 1, No. 2, 2011, pp. 147–161. doi: 10.1615/IntJUncertaintyQuantification.v1.i2.30

[72] Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S., *Global Sensitivity Analysis: The Primer*, 1st ed. Wiley-Interscience: Hoboken, NJ, 2008.

[73] Ratto, M., Pagano, A., and Young, P., "State Dependent Parameter metamodelling and sensitivity analysis," *Computer Physics Communications*, Vol. 177, No. 11, 2007, pp. 863–876. doi: 10.1016/j.cpc.2007.07.011

[74] Cukier, R. I., Fortuin, C. M., Shuler, K. E., Petschek, A. G., and Schaibly, J. H., "Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I Theory," *Journal of Chemical Physics*, Vol. 59, No. 8, 1973, pp. 3873–3878. doi: 10.1063/1.1680571

[75] Lamorte, N., Glaz, B., Friedmann, P. P., Culler, A. J., Crowell, A. R., and Mcnamara, J. J., "Uncertainty Propagation in Hypersonic Aerothermoelastic Analysis," *51st AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials and Co-located Conferences*, AIAA 2010-2964, Apr. 2010.

[76] Lamorte, N., "Uncertainty Propagation in Hypersonic Vehicle Aerothermoelastic Analysis," 2013.

[77] Lamorte, N., and Friedmann, P. P., "Uncertainty Propagation in Integrated Airframe-Propulsion System Analysis for Hypersonic Vehicles," *17th AIAA International Space Planes and Hypersonic Systems and Technologies Conference*, AIAA 2011-2394, 2011.

[78] Hosder, S., "Efficient Uncertainty Quantification & Sensitivity Analysis for Hypersonic Flow and Material Response Simulations Under Inherent and Model-Form Uncertainties," *5th Ablation Workshop*, 2012.

[79] Hu, Z., and Mahadevan, S., "Global sensitivity analysis-enhanced surrogate (GSAS) modeling for reliability analysis," *Structural and Multidisciplinary Optimization*, Vol. 53, No. 3, 2016, pp. 501–521. doi: 10.1007/s00158-015-1347-4

[80] Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., and Tarantola, S., "Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index," *Computer Physics Communications*, Vol. 181, No. 2, 2010, pp. 259–270. doi: 10.1016/j.cpc.2009.09.018

[81] Sobol, I. M., "Quasi-Monte Carlo methods," *Progress in Nuclear Energy*, Vol. 24, No. 1–3, Jan. 1990, pp. 55–61. doi: 10.1016/0149-1970(90)90022-W

[82] Halton, J. H., "Algorithm 247: Radical-inverse quasi-random point sequence," *Communications of the ACM*, Vol. 7, No. June, 1964, pp. 701–702. doi: 10.1145/355588.365104

[83] Hammersley, J. M., and Handscomb, D. C., *Monte Carlo Methods*. London Chapman and Hall, 1965.

[84] Aihara, M., Takami, H., Sawada, Y., Morimoto, S., Kariya, K., Kudo, I., Ueno, K., Kimura, A., and Yoshida, Y., "Effect of fibronectin and von willebrand factor on the adhesion of human fixed washed platelets to collagen immobilized beads," *Thrombosis Research*, Vol. 44, No. 5, 1986, pp. 661–672

[85] Akaike, H., "A New Look at the Statistical Model Identification," *IEEE Transactions on Automatic Control*, Vol. 19, No. 6, 1974, pp. 716–723. doi: 10.1109/TAC.1974.1100705

[86] Nuggehally, M. A., Shephard, M. S., Picu, C. R., and Fish, J., "Adaptive model selection procedure for concurrent multiscale problems," *International Journal for Multiscale Computational Engineering*, Vol. 5, No. 5, 2007, pp. 369–386. doi: 10.1615/IntJMultCompEng.v5.i5.20

[87] Hombal, V. K., and Mahadevan, S., "Model Selection Among Physics-Based Models," *Journal of Mechanical Design*, Vol. 135, No. 2, Jan. 2013, pp. 21003–21015.

[88] Murphy, K. P., "Conjugate Bayesian analysis of the Gaussian distribution," 2007. [Online]. Available: http://www.cs.ubc.ca/~murphyk/Papers/bayesGauss.pdf