공학박사학위논문

최적화 기반 모델 보정을 위한 순차적 최적화 및 불확실성 확산 기법

Sequential Optimization and Uncertainty Propagation for Optimization-Based Model Calibration

2019 년 2 월

서울대학교 대학원

기계항공공학부

이 규 석

Abstract

Sequential Optimization and Uncertainty Propagation for Optimization-Based Model Calibration

Guesuk Lee
Department of Mechanical and Aerospace Engineering
The Graduate School
Seoul National University

Highly credible computational models have long been a dream of engineers. One factor that impacts the credibility of a computational model is the existence of unknown input variables in the model. For this reason, model calibration – a process of estimating unknown input variables in a computational model – has been explored, with the goal of providing solutions that could ultimately improve the credibility of computational models. Optimization-based model calibration (OBMC) is recognized as a promising solution for estimating the unknown input variables in a computational model through the use of optimization techniques. For OBMC, a question of fundamental importance arises: How can OBMC be carried out accurately and efficiently under various sources of uncertainties and errors? In order to facilitate OBMC for model calibration in a statistical sense, this doctoral dissertation aims to address four essential issues: 1)

Research Thrust 1 – Characterize the uncertainty in experimental observations considering the systematic and random measurement errors; 2) Research Thrust 2 – Derive analytical sensitivity information for checking the convexity of the optimization problem formulated by OBMC, and conduct robust OBMC using the derived analytical sensitivity information; 3) Research Thrust 3 – Formulate an optimization under uncertainty loop for accurate and efficient OBMC, and the associated optimization and uncertainty propagation processes; 4) Research Thrust 4 – Validate the calibrated computational model that is derived from OBMC not only in a statistical sense, but also with a straightforward explanation.

Research Thrust 1: The process of model calibration estimates unknown input variables of a computational model with a goal of maximizing the agreement (or minimizing the disagreement) between probability distributions that result from computational predictions and experimental observations. To execute accurate model calibration, a proper probability distribution is required that describes the uncertainty in the experimental observations (data). However, experimental observations may include systematic and random measurement errors. When characterizing the uncertainty in the experimental observations, no consideration of systematic and random measurement errors may degrade the accuracy of calibrated results. Thus, Research Thrust 1 proposes a method that utilizes maximum likelihood estimation with modeling of systematic and random errors to properly develop a probability distribution that describes the uncertainty in the experimental observations.

Research Thrust 2: Occasionally, gradient-based optimization algorithms are effective for use in OBMC. However, the calibrated results derived from gradient-based algorithms that use existing calibration metrics result in inaccurate and unstable calibration. Therefore, Research Thrust 2 aims to 1) investigate the fundamental explanations of the inaccurate and unstable calibrated results that arise from using existing calibration metrics, and 2) enhance the robustness of OBMC by providing gradient information.

Research Thrust 3: OBMC is a probabilistic method used to estimate unknown input variables through the use of optimization under uncertainty (OUU). OUU combines the optimization process with the probabilistic analysis; this is used for the uncertainty propagation process in OBMC. Performing OBMC using an OUU formulation requires a high computational cost because the optimization and uncertainty propagation processes are associated in a loop. To improve the efficiency of OBMC, Research Thrust 3 presents a sequential OBMC approach that makes use of first 1) an efficient, and then 2) a highly accurate uncertainty propagation method, in sequence. The proposed method is called sequential optimization and uncertainty propagation (SOUP).

Research Thrust 4: As the final process of model calibration, model validation checks whether the calibrated result is valid or not. The validation should be conducted in a quantitative and statistical way. Research Thrust 4 proposes a new validation metric called probability of coincidence (POC). The POC calculates the probabilistic degree to which the computational prediction agrees with the experimental observations.

Keywords: Optimization-Based Model Calibration (OBMC)
Statistical Model Calibration and Validation
Uncertainty Characterization
Systematic and Random Measurement Errors
Optimization Under Uncertainty (OUU)
Sequential Optimization and Uncertainty Propagation (SOUP)
Probability of Coincidence (POC)

Student Number: 2014-21865

Table of Contents

List of Tables

List of Figures

Nomenclature

OBMC	Optimization-based model calibration
UP	Uncertainty propagation
RBDO	Reliability-based design optimization
OUU	Optimization under uncertainty
SOUP	Sequential optimization and uncertainty propagation
QOI	Quantity of interest
PDF	Probability density function
CDF	Cumulative density function
MCS	Monte Carlo simulations
UDR	Univariate dimension reduction
MMM	Moment matching metric
V&V	Verification and validation
POC	Probability of coincidence
POS	Probability of separability
TFT-LCD	Thin-film transistor liquid crystal display
$\mathbf{X}$	A vector of input variables in a computational model
$\mathbf{X}_{\mathrm{known}}/\mathbf{X}_{\mathrm{k}}$	A vector of known input variables in a computational model
$\mathbf{X}_{\mathrm{unknown}}/\mathbf{X}_{\mathrm{u}}$	A vector of unknown input variables in a computational model
$x_{\mathrm{unknown}}/x_{\mathrm{u}}$	An unknown input variable in a computational model
$\boldsymbol{\theta}$	A vector of statistical parameters/A test statistic of hypothesis testing
$\boldsymbol{\theta}_{\mathbf{X}\mathrm{unknown}}/\boldsymbol{\theta}_{\mathbf{X}\mathrm{u}}$	A vector of statistical parameters of unknown input variables (calibration variables of OBMC)
$\boldsymbol{\theta}_{\mathbf{X}\mathrm{u},\, k=1}$	The initial calibration parameters of OBMC (k = 1)

$\hat{\boldsymbol{\theta}}_{\mathbf{X}\mathrm{u}}$	The intermediate calibration parameters of OBMC by SOUP (the final calibration parameters of OBMC by the 1st sequence of SOUP)
$\boldsymbol{\theta}_{\mathbf{X}\mathrm{u,\,opt}}/\hat{\boldsymbol{\theta}}_{\mathbf{X}\mathrm{u,\,opt}}$	The final (optimal) calibration parameters of OBMC/The final calibration parameters of OBMC by the 2nd sequence of SOUP)
f	A calibration metric (an objective function for OBMC)
$\mathbf{Y}$	A vector of system responses
y	A random variable for system response
$\mathbf{Y}_{\mathrm{obs}}$	A set (vector) of observed system responses
y_{obs}	A datum from a set of observed system responses
$\mathbf{Y}_{\mathrm{pre}}$	A vector of predicted system responses
y_{true}/y	The random variable of the true variability for system responses
$p(\cdot)$	A probability density function
α	The maximum amount of systematic measurement error (bias)
n	The number of experimental observations
e	The measurement error
e_{sys}	The systematic measurement error
e_{ran}	The random measurement error
μ_y	The mean of the system responses
σ_y	The standard deviation of the system responses
s_y	The variance of the system responses
$\mu_{\mathbf{Y}\mathrm{obs}}$	The mean of observed system responses
$\mu_{\mathbf{Y}\mathrm{pre}}$	The mean of predicted system responses
L	The likelihood function
$\mathbf{x}$	A vector of variables for a function
$\mathbf{x}^*$	A vector of the given point for Taylor's expansion
$\mathbf{H}$	The Hessian matrix

R	The remainder term of Taylor's expansion
$f(\mathbf{x})$	The function of N-variables
Δf	The change in a function
$\det(\cdot)$	The determinant of a matrix
A	A set of the variable
a_i	Nth input variable of a linear system model ($i = 1, \ldots, N$)
N	The number of input variables in a model
b	The remaining input variable of a linear system model
PR	The probability residual function
PR_{abs}	The probability residual function by the integral of the absolute difference
PR_{squ}	The probability residual function by the integral of the squared difference
$\triangledown$	The vector of the first-order derivatives
$\mathbf{H}$	The Hessian matrix including the second-order derivatives
k	The iteration number of an optimization problem
$\boldsymbol{x}_k$	A vector of design variables for an optimization problem in the kth iteration
$\mathbf{p}_k$	A vector for search direction of an optimization problem in the kth iteration
α_k	The step length for an optimization problem in the kth iteration
MMM	The moment matching metric function
MMM_{abs}	The moment matching metric function by the summation of the absolute difference
MMM_{squ}	The moment matching metric function by the summation of the squared difference
M_{model}	A computational model with N number of input variables: $y = M_{\text{model}}(x_1, x_2, \ldots, x_N)$
m	The iteration number for sampling methods

$E(\cdot)$	The expectation operator
$I(\cdot)$	The multivariate integration
y_{a}	The univariate approximation
d_{area}	The area metric function
$F(y)$	The cumulative density function of predicted system responses
$S_n(y)$	The empirical cumulative density function of *n* numbers of experimental observations
POC	The probability of coincidence function
H_0	The null hypothesis
$\mathrm{H}_1/\mathrm{H}_{\mathrm{a}}$	The alternative hypothesis
p_{f}	The probability of failure
L	A random variable of the load
S	A random variable of the strength
P_{NS}	The probability of a non-separable region
POS	The probability of separability function
P_{OV}	The probability of overlapped area
$\tilde{x}$	The median value
y_{D}	The cantilever beam's deflection at the free end of the beam
$y_{\mathrm{D,\ obs}}$	The observed cantilever beam's deflection at the free end of the beam
$\mathbf{Y}_{\mathrm{D,\ obs}}$	The observed data set of the cantilever beam's deflection
P_{D}	The vertical downward loading to a cantilever beam
q_{u}	The bearing capacity
c_{s}	The cohesion of soil
γ_{s}	The unit weight of soil
$B_{\mathrm{s}}/H_{\mathrm{s}}$	The width and height of a strip footing
$N_{\mathrm{c}}/N_{\mathrm{q}}/N_{\gamma}$	The bearing capacity factors from the given information on the angle of internal friction of the soil

M_i	The ith statistical moment (e.g., the first statistical moment (M_1): mean)
$\rho_{\mathrm{wheel,\ alum}}$	The density of aluminum in the wheel part
$t_{\mathrm{wheel,\ poly}}$	The thickness of polyurethane in the wheel cover part
k_{air}	The stiffness of spring element between the steering wheel and air bag
E_{g}	The Young's modulus of glass
E_{py}	The Young's modulus of polarizer

Chapter 1

Introduction

1.1 Motivation

Today, engineering is a discipline that requires significant use of computational and software instruments. Computer-assisted design, computational fluid dynamics, and finite-element analysis applications are some of the basic instruments that engineers deploy when creating or improving new engineered products or systems (Molina et al. 1995) (Benek et al. 1998) (Zhan et al. 2011) (Xu et al. 2011) (Shi et al. 2012) (Fender et al. 2014) (Lee and Gard 2014) (Silva and Ghisi 2014) (Zhu et al. 2016) (Jung et al. 2016). Although various computer-aided instruments have enhanced the power of engineers, the credibility of the use of computational models in real-world applications has been a growing concern. To this end, this concern has inspired significant research interest in using model calibration and validation techniques to enable more accurate and trustworthy computational models (Babuska and Oden 2004) (Hills et al. 2008) (Oberkampf and Trucano 2008) (Kwaśniewski 2009) (Roy and Oberkampf 2010) (Sargent 2013) (Oden et al. 2013) (Mousaviraad et al. 2013) (Borg et al. 2014) (Sankararaman and Mahadevan 2015).

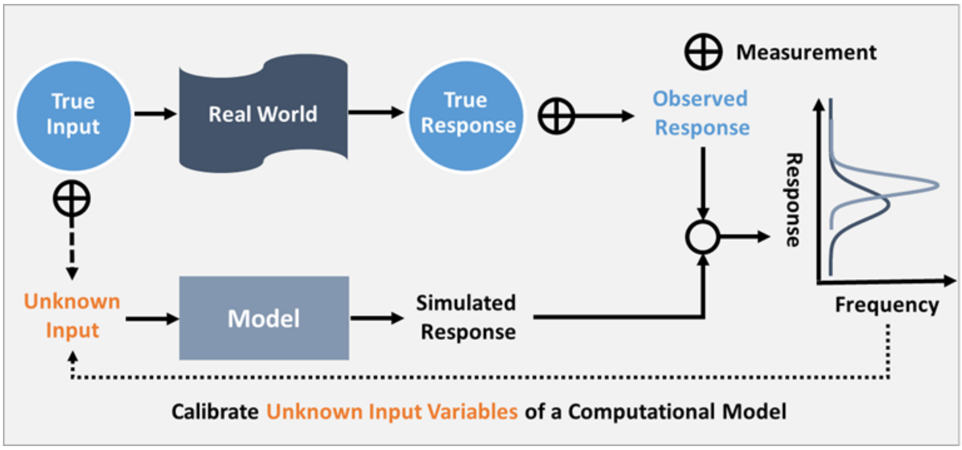

Figure 1-1 Scheme of Statistical Calibration of the Unknown Input Variables of a Computational Model

The existence of unknown input variables in a computational model is one reason for the low predictive capability of the model. In many cases, establishing and conducting physical experiments to measure quantities of the unknown input variables is time consuming and expensive. In contrast, conducting experiments to observe system responses is often more feasible. Therefore, efforts have been made to devise model calibration to estimate the unknown input variables by comparing the system responses from computational predictions and experimental observations, as shown in Figure 1-1 (Kennedy and O'Hagan 2001) (Campbell 2006) (Higdon et al. 2008) (Youn et al. 2011) (Zhan et al. 2011) (Arendt et al. 2012b).

Despite the increasing interest in model calibration, it remains a difficult task due to two important challenges (Kennedy and O'Hagan 2001) (Campbell 2006). The first challenge is that model calibration requires solving an implicit inverse function. For most computational models that emulate the behavior of an engineered system, it is almost impossible to obtain a closed form (explicit form) of the inverse function. Thus, an analytical solution for model calibration in real-world applications is not available. The second challenge is that model calibration must address issues related to the uncertainties in the experiments and computational models (Oberkampf et al. 2002) (Chen et al. 2004) (Helton and Davis 2003) (Ferson et al. 2004) (Bayarri et al. 2007). Due to the various sources of uncertainties, deterministic calibration of a set of unknown input variables can significantly degrade the predictive capability of a computational model. Therefore, statistical approaches are required to statistically calibrate the unknown input variables. To deal with the challenges outlined above, optimization-based model calibration has been recognized as a promising solution to statistically calibrate the unknown input variables using optimization techniques (Youn et al. 2011) (Zhan et al. 2011) (Jung et al. 2016) (Lee et al. 2018). Despite the increasing interest in calibrating unknown input variables using optimization techniques, a trustworthy optimization-based model calibration method will only be available for use after several important technical issues are addressed. Section 1.2 overviews optimization-based model calibration and introduces four technical issues.

1.2 Research Scope and Overview

The objective of optimization-based model calibration (OBMC) is to estimate the statistical parameters ($\boldsymbol{\theta}$) of unknown input variables ($\mathbf{X}_{\text{unknown}}$) in a computational model by formulating an optimization problem. The following equation presents the mathematical formulation of OBMC that uses an unconstrained optimization problem (Lee et al. 2018):

$$\underset{\boldsymbol{\theta}_{\mathbf{X}_{\text{unknown}}}}{\text{Maximize}} \text{ or } \underset{\boldsymbol{\theta}_{\mathbf{X}_{\text{unknown}}}}{\text{Minimize}} f\left(p_{\mathrm{Y}}(\mathbf{Y}_{\text{obs}}), p_{\mathrm{Y}}\left(\mathbf{Y}_{\text{pre}}\left(\boldsymbol{\theta}_{\mathbf{X}_{\text{known}}}, \boldsymbol{\theta}_{\mathbf{X}_{\text{unknown}}}\right)\right)\right) \quad (1\text{-}1)$$

where $\boldsymbol{\theta}_{\mathbf{X}_{\text{unknown}}}$ is a set of calibration parameters (design variables of an optimization problem) and f denotes a calibration metric (an objective function of an optimization problem). (The input variables ($\mathbf{X}$) of a computational model can be deterministic. In this case, only the mean value is incorporated as a deterministic input variable.) In a probabilistic sense, the observed ($\mathbf{Y}_{\text{obs}}$) and predicted ($\mathbf{Y}_{\text{pre}}$) system responses are characterized by probability distributions ($p(\cdot)$). A calibration metric (f) quantifies the agreement or disagreement between the two probability distributions that arise from the observations and predictions. By maximizing or minimizing the calibration metric, the statistical parameters ($\boldsymbol{\theta}$) of the unknown input variables ($\mathbf{X}_{\text{Unknown}}$) are calibrated toward making computational predictions ($\mathbf{Y}_{\text{pre}}$) that are in accordance with experimental observations ($\mathbf{Y}_{\text{obs}}$).

Figure 1-2 presents the overall scheme of OBMC. OBMC consists of four important processes: 1) uncertainty characterization of experimental observations; 2) iterative optimization until the calibration metric (f), which is the objective function of OBMC, converges to its maximum or minimum; 3) uncertainty propagation to obtain the probability distribution ($p(\mathbf{Y}_{\text{pre}})$) of the predicted system responses; 4) a validity check of the calibrated results. For each process, technical issues should be addressed to improve the credibility of the computational models that are derived from OBMC.

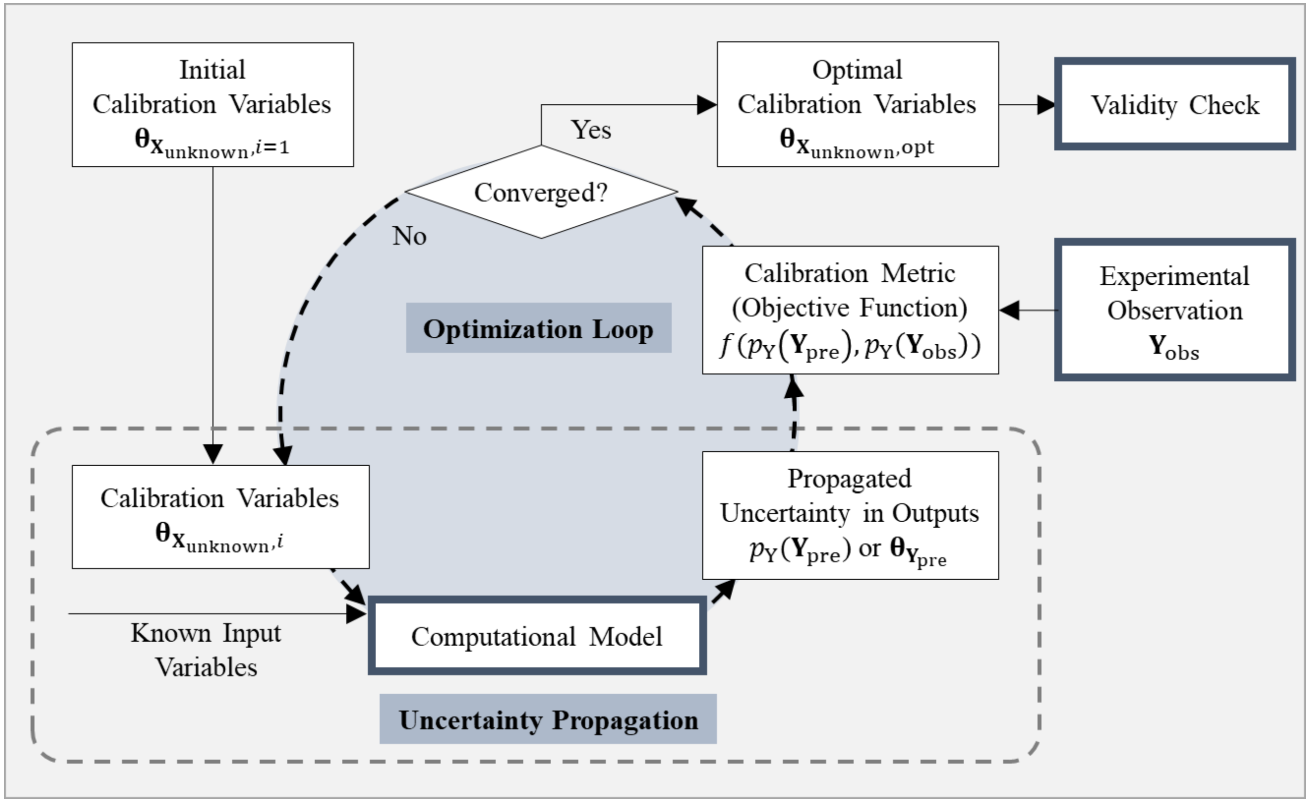

Figure 1-2 Overall Scheme of Optimization-Based Model Calibration

This doctoral dissertation aims to advance OBMC by tackling four technical issues: 1) the existence of systematic and random measurement errors in experimental observations; 2) the inaccuracy and instability of the calibrated results derived from OBMC; 3) the inefficiency of optimization under uncertainty for OBMC; and 4) a procedure for a validity check of the calibrated results. To address the above-mentioned technical issues, the research scope described in this doctoral dissertation is to develop technical advances in the following four research thrusts:

Research Thrust 1: Uncertainty Characterization of Experimental Observations Considering Systematic and Random Measurement Errors Using Maximum Likelihood Estimation

Uncertainty characterization, also called uncertainty modeling, is the science of quantitative characterization of uncertainties in engineering applications (Parry 1996) (Agarwal et al. 2004) (Helton et al. 2006) (Ghanem et al. 2008) (McFarland and Mahadevan 2008a) (Jung et al. 2011) (XI et al. 2013). One of the main activities in uncertainty characterization is to characterize the uncertainties in given data. As previously explained, OBMC requires that a description of the uncertainties in experimental observations be compared with the predicted system responses (Figure 1-3). Various sources of uncertainty, which include physical uncertainties in geometry and material properties, and systematic and random measurement errors, affect the variability observed in experimental observations. Among them, systematic and random measurement errors are often disregarded in the uncertainty

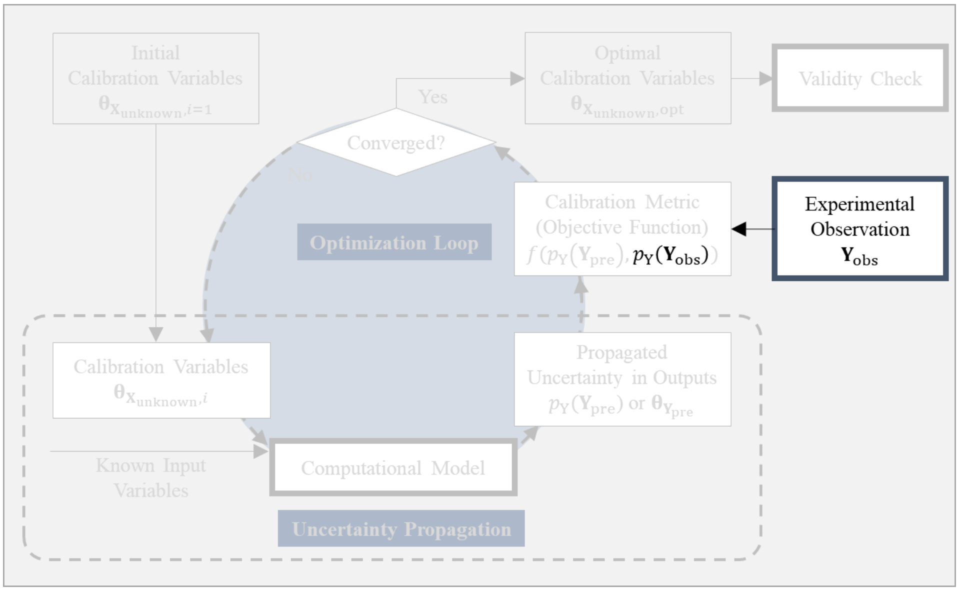

Figure 1-3 Uncertainty Characterization of Experimental Observations in Optimization-Based Model Calibration

characterization process, even though they may be responsible for much of the variability in the experimental observations. To address this issue, Research Thrust 1 proposes an uncertainty characterization method to consider systematic and random measurement errors. The proposed method separately distinguishes each source of measurement error by using a specific type of probability distribution. Then, statistical parameters of each assumed probability distribution are estimated by adopting the maximum likelihood estimation.

Research Thrust 2: Robust Optimization-Based Model Calibration with Analytical Sensitivity Information

OBMC adopts an optimization algorithm to calibrate the unknown input variables of a computational model (Youn et al. 2011) (Zhan et al. 2011) (Jung et al. 2016) (Lee et al. 2018). As an important element of OBMC, a calibration metric is defined as an objective function that can quantify the difference or the similarity between the system responses derived from computational predictions and experimental observations (Figure 1-4). However, OBMC utilizing existing calibration metrics (e.g., likelihood function, probability residual) has resulted in inaccurate and unstable model calibrations. For example, 1) (inaccurate calibration) the calibrated result may converge to large standard deviations or biased mean values, as shown in Figure 1-5c and Figure 1-5d; 2) (instable calibration) the optimization may diverge rather than converge. To work to address this issue, Research Thrust 2 investigates the possible reasons (e.g., non-convexity of the optimization problem, local minima, and others) to address the inaccuracy and instability of the calibrated results derived from OBMC. As part of the research, analytical sensitivity information (the first- and second order derivatives of the calibration metrics (objective functions) with respect to calibration parameters (statistical parameters of unknown input variables) is derived with assumptions. Ultimately, it is proved that OBMC performed using existing calibration metrics is not globally convex, and OBMC performed with analytical sensitivity information showed better accuracy and stability.

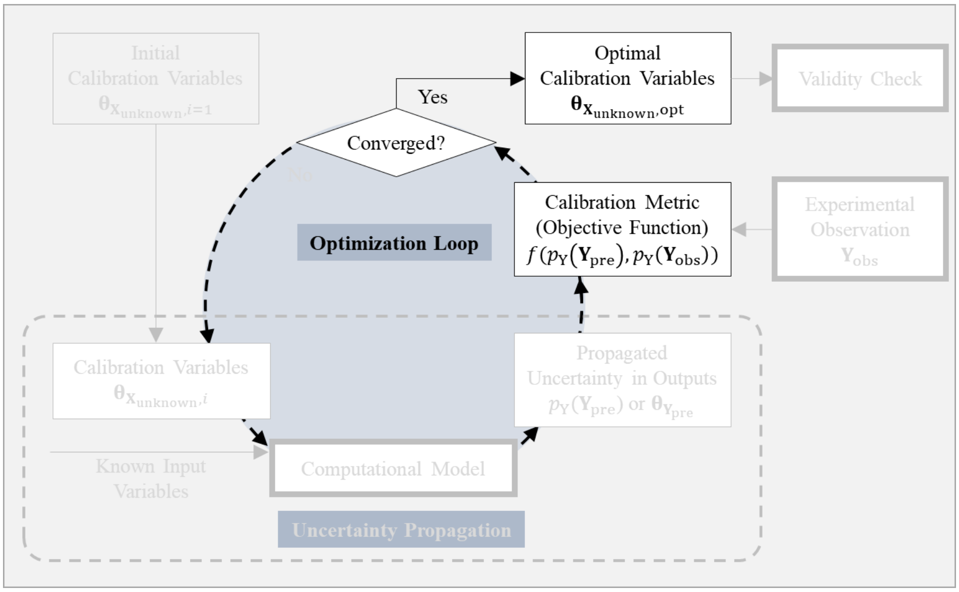

Figure 1-4 Iterative Optimization of Optimization-Based Model Calibration with a Calibration Metric (an Objective Function for Optimization-Based Model Calibration)

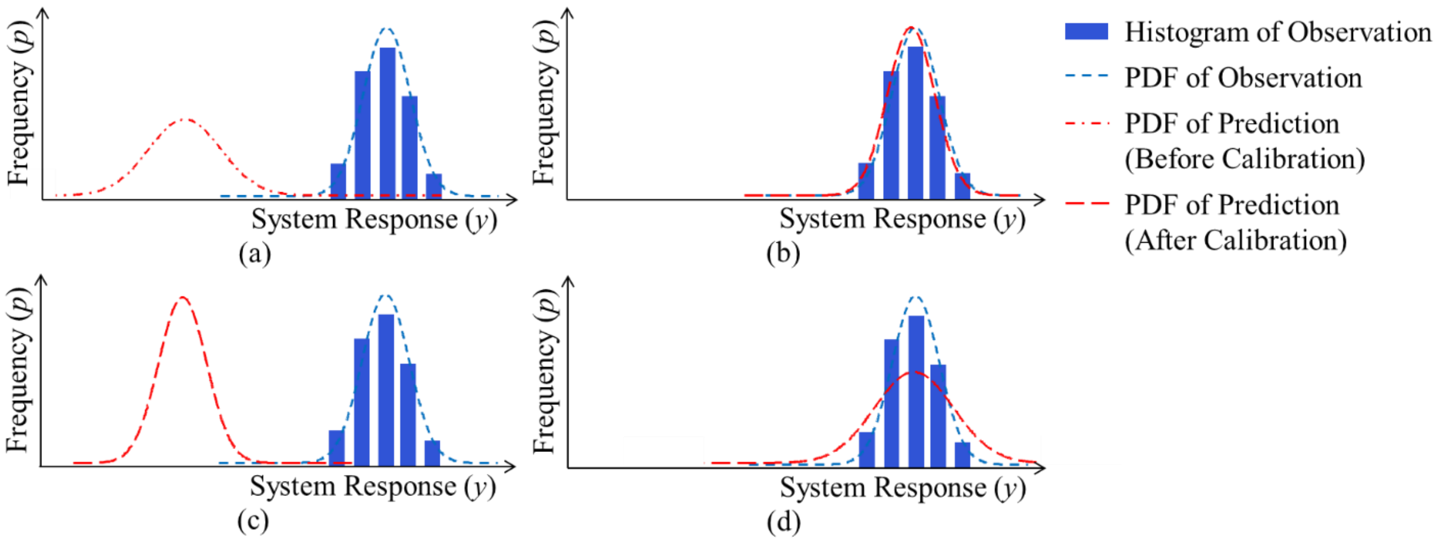

Figure 1-5 Illustration of Optimization-Based Model Calibration Results: (a) Two PDFs of Observations and Predictions Before Calibration, (b) an Accurate Calibrated Result, (c) and (d) Inaccurate Calibrated Results (Lee et al. 2018)

Research Thrust 3: Sequential Optimization and Uncertainty Propagation for Optimization-Based Model Calibration

OBMC is a probabilistic way to estimate statistical parameters (e.g., the mean and standard deviation) of the unknown input variables through the use of optimization techniques. Performing optimization in a probabilistic sense adopts optimization under uncertainty (OUU), which requires a high computational cost (Eldred et al. 2002) (Agarwal et al. 2004) (Liu et al. 2006) (Yao et al. 2011). OBMC that implements OUU associates the optimization process with the uncertainty

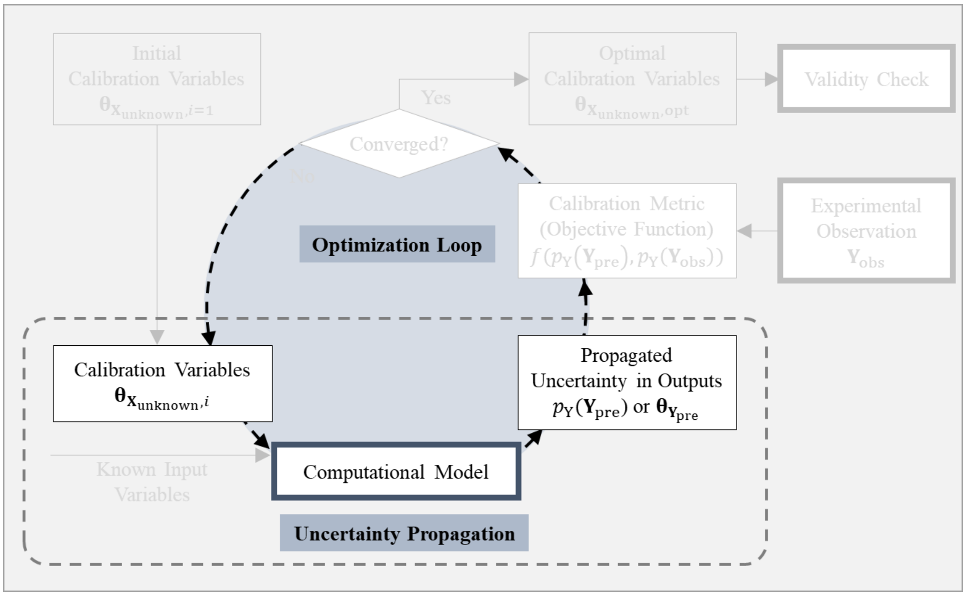

Figure 1-6 Using Optimization Under Uncertainty to Associate the Optimization and Uncertainty Propagation Processes for Optimization-Based Model Calibration

propagation (UP) process (Figure 1-6), also called probabilistic assessment or probabilistic analysis, where the uncertainties in the input variables are propagated through a model to obtain the variability in the outputs or system responses of a system. Since the optimization and UP processes are both computationally expensive, it costs even more when the two processes are associated in a loop. Therefore, increasing the efficiency of OBMC, while retaining high accuracy, is an important issue deserving research attention. To improve the efficiency of OUU for OBMC, Research Thrust 3 proposes a sequential optimization-based model calibration approach. The proposed idea is based on a comprehensive review of OUU formulations and uncertainty propagation methods from the society of reliability-based design optimization, one category of OUU.

Research Thrust 4: Validation Metric – Probability of Coincidence

Model validation is the process of determining the degree to which a computational model is an accurate representation of the real phenomenon, from the perspective of the model's intended uses (Babuska and Oden 2004) (Hills et al. 2008) (Oberkampf and Trucano 2008) (Weathers et al. 2009) (Sankararaman and Mahadevan 2011) (Ling and Mahadevan 2013). Model validation can be executed at the completion of model calibration to check whether the calibrated result is valid or not (Figure 1-7). Many research efforts have been made to develop statistical model validation in order to quantitatively and statistically determine the degree of validity (Oberkampf and Barone 2006) (Ferson et al. 2008) (Liu et al. 2011). Among them, the area metric possesses most of the desirable features of a validation metric for quantitative comparison of experiments and simulations. Hypothesis testing could be devised to evaluate the null hypothesis – if a calibrated model is valid – by comparing the computed area metric value with a designated critical value. However, there is still a great need for a new validation metric. First, a new validation metric should be able to be calculated with a full description (probability distribution) of the experimental observations from the validation experiments. Second, the results of statistical validation should be straightforwardly probabilistic; thereby, an analyst can make decisions based on the results of the validity check. To address these needs, Research Thrust 4 proposes a new validation metric, called probability of coincidence.

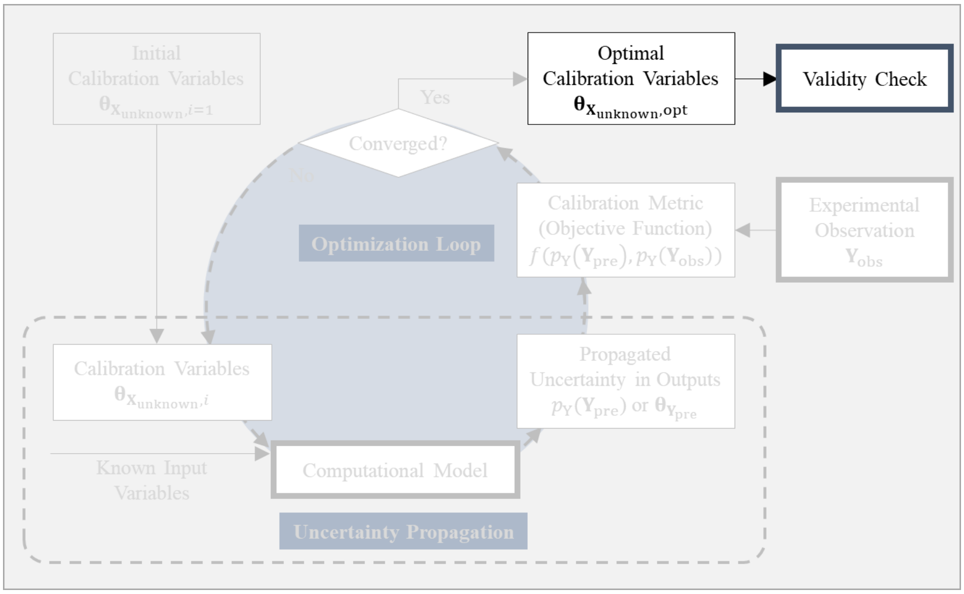

Figure 1-7 A Validity Check of the Calibrated Results by Optimization-Based Model Calibration

1.3 Dissertation Layout

This doctoral dissertation is organized as follows. Chapter 2 proposes an uncertainty characterization method that considers systematic and random errors (Research Thrust 1). Chapter 3 presents a method to enhance robustness of OBMC by providing analytical sensitivity information (Research Thrust 2). Chapter 4 proposes a sequential optimization and uncertainty propagation method for efficient OBMC (Research Thrust 3). Chapter 5 proposes a new validation metric – probability of coincidence – to check the validity of the calibrated results (Research Thrust 4). In Chapter 6, four engineering problems (uncertainty characterization of observed cantilever beam deflections, and model calibration of a bearing capacity equation, an automobile steering column and wheel vibrational model, and a thin-film transistor liquid crystal display deflection model) are employed to demonstrate the proposed research thrusts. Finally, Chapter 7 discusses the contributions of this dissertation and potential future research directions.

Sections of this chapter have been published or submitted as the following journal article:

1) Guesuk Lee, Wongon Kim, Hyunseok Oh, and Byeng D. Youn, Nam H. Kim "Review of Statistical Model Calibration and Validation – From the Perspective of Uncertainty Structures," *Structural and Multidisciplinary Optimization*, Submitted in September 2018.

Chapter 2

Uncertainty Characterization of Experimental Observations Considering Systematic and Random Measurement Errors

The true value of the quantity of interest (QOI) is usually estimated using measurements or observations gathered by conducting physical experiments. (For conducting statistical model calibration, characterizing the uncertainty in measured system responses (the QOI) must be compared with the uncertainty in the predicted system responses through the use of a computational model.) To improve the estimate of the true value (or the true uncertainty characterized by a probability distribution) of the QOI, repeated measurements in the same experimental conditions are often conducted. Multiple measurements are used to enable consideration of the various sources of uncertainties, such as physical uncertainties or measurement errors. Thus, the variability due to physical or inherent uncertainties (e.g., geometric uncertainty – manufacturing tolerance, material uncertainty – elastic modulus of a material) becomes the subject to be accurately characterized. On the other hand, measurement errors are a barrier to accurate estimation of the true value of the QOI.

Measurement error, or observational error, refers to the difference between the measured value of a quantity and its true value. Carrying out a perfect, error-free experiment is impossible. Whether its degree is large or small, a measured value cannot avoid measurement errors. Furthermore, in many cases, physical experiments are carried out across different experimental conditions. In this case, experimental observations from different experimental conditions should be aggregated to minimize the degree of statistical uncertainty present due to an insufficient amount of data.

Significant efforts have been made to quantify the true variability in a QOI. To the best of the author's knowledge, most previous works have focused on examining how to quantify the variability with a limited amount of observed data; these efforts have ignored measurement errors. Other prior studies have been conducted with a focus on eliminating sources of measurement errors, for example, making effort to 1) maintain the same experimental conditions, 2) continuously calibrate measurement instruments, or 3) use expensive instruments to minimize random errors in the measured values. In many cases, however, it may not be possible to invest a large expense in experiments. Even when funding is available, results from expensive equipment still can contain errors.

Chapter 2 thus proposes an uncertainty characterization method that considers measurement errors. Especially for model calibration, which estimates statistical parameters of unknown input variables in a computational model by comparing probability distributions that describe the uncertainties in observations and predictions, accurate calibration requires accurate characterization or modeling of the uncertainty in experimental observations by considering measurement errors in the data. The proposed method offers two primary contributions: 1) Accurate characterization of the true variability in a QOI; 2) Additional data that can be derived under different experimental conditions.

The remainder of Chapter 2 is organized as follows. Section 2.1 provides an overview of measurement errors that can be observed in engineering problems. Section 2.2 explains how the proposed method is formulated to develop a probability distribution for the true variability in a QOI, considering measurement errors. The conclusions for Research Thrust 1 are provided in Section 2.3. Case studies for proving the effectiveness of the proposed method are presented in Chapter 6.

2.1 Measurement Errors Observed in Engineering Problems

Section 2.1 presents the research topics of Research Thrust 1. Section 2.1.1 includes a brief discussion of uncertainty and variability to aid in understanding measurement errors in engineering problems. Section 2.1.2 explains measurement errors, which include systematic and random measurement errors. Section 2.1.3 provides a discussion on how to model measurement errors using parametric probability distributions.

2.1.1 Uncertainty and Variability

When it is said that there is “uncertainty” in a QOI, it means a lack of certainty or a state of limited knowledge that prevents exact determination of a certain value of a

QOI. Thus, in general, uncertainty in a QOI is characterized or quantified by a probability distribution, which is assigned based upon the information or evidence about the likelihood of what the true value might be (Soundappan et al. 2004) (Guo and Du 2007) (Lin et al. 2014). The engineering information for a QOI refers, for example, to the values of data, the number of data, the type of probability distribution, the upper and lower bound, and other related information.

In measured or observed data of a QOI, "variability" exists due to physical uncertainties and measurement errors (Zhang and Mahadevan 2000) (Oberkampf et al. 2004a) (Buranathiti et al. 2006) (Hills 2006) (Xie et al. 2007) (Urbina et al. 2011) (Sankararaman et al. 2011) (Zhang et al. 2013) (Jung et al. 2014) (Zhu et al. 2016). Physical uncertainty arises due to the natural inherent uncertainty in material and geometric properties (e.g., inherent uncertainty in the elastic modulus and manufacturing tolerances). In general, physical uncertainty is considered aleatory uncertainty, which means irreducible uncertainty. In this case, the goal is to quantify the uncertainty well using a probability distribution in the uncertainty characterization process. However, one challenge in accurately characterizing the variability in observed data by physical uncertainties is the existence of measurement errors.

2.1.2 Systematic and Random Measurement Error

Measurement errors, or observational errors, are another cause of difference between measured values of a QOI and its true value (Hills 2006) (Harmel et al. 2010) (da Silva Hack and Schwengber ten Caten 2012) (Ling and Mahadevan 2013) (Uusitalo et al. 2015). The true value denotes the variability in a QOI that results only from the previously introduced physical uncertainty sources. Measurement errors can be categorized into two types: 1) systematic measurement error, and 2) random measurement error (Ferson and Ginzburg 1996) (Easterling 2001) (Liang and Mahadevan 2011) (Ling and Mahadevan 2013).

Systematic measurement error, also known as measurement bias, is introduced by inaccuracy factors that occur during the observation or measurement process. For example, when a particular sensor is used in all of the replicate tests, or when all replicate tests are conducted in a certain experimental setting (e.g., a higher temperature than a normal condition), then all measurements may have a similar biased error. For another example, a system response, which is the QOI for conducting model calibration, can be measured under different experimental conditions. From the standpoint of one experimental condition, other experimental conditions can be considered to be systematic measurement errors. In this case, the observations from two different experimental conditions can be integrated into one data set, which can be used for estimating the unknown input variables.

Random measurement error is caused by poor precision factors. Human errors, like fluctuations in the experimenter's interpretation of the instrumental readings, are one example. As another example, measurements may be gathered from a sensor that is picked randomly from a population of sensors; in this case, multiple measurements may have inaccuracy that is described by a distribution of error.

Figure 2-1 helps to explain systematic and random measurement errors. Each data set of the green, blue, and red shaded part in Figure 2-1 denotes ten repeated experiments. In Figure 2-1a, the data in the upper part (the green shaded part) describe the true variability, or the value of the QOI, which results only from the previously introduced physical uncertainty sources. If there is no measurement error in the experiments while the inherent uncertainties exist, the observed data should appear like the data in the upper part of Figure 2-1a. The green probability distribution of the right-hand side exhibits the mother population of the QOI; this can be quantified or characterized after a perfect uncertainty characterization process. On the other hand, the sample data in the lower part (the blue shaded part) of Figure 2-1a describe that each datum is biased due to systematic error. Note that each datum has not been biased with the same magnitude of error – this is an important clarification. The blue probability distribution presents that the mother population of the QOI (the above red probability distribution) is biased. In Figure 2-1b, the data (the red shaded part) describe the observations that result from a situation where both systematic and random errors are present. For example, in the third observation from the left, the datum (the upper red point) is observed to deviate slightly from the datum (the lower blue point) that exhibits only systematic error. (Note that random measurement error does not affect the average, only the variability around the average.) In the presence of systematic and random measurement errors, the observed data will be like the data shown in Figure 2-1b. If the systematic and random measurement exist, the probability distribution will be wrongly characterized as the red probability distribution. The goal of this research is to characterize the variability in the true value (the green probability distribution in Figure 2-1a) by using the observed data and subtracting the influence by the systematic and random measurement errors.

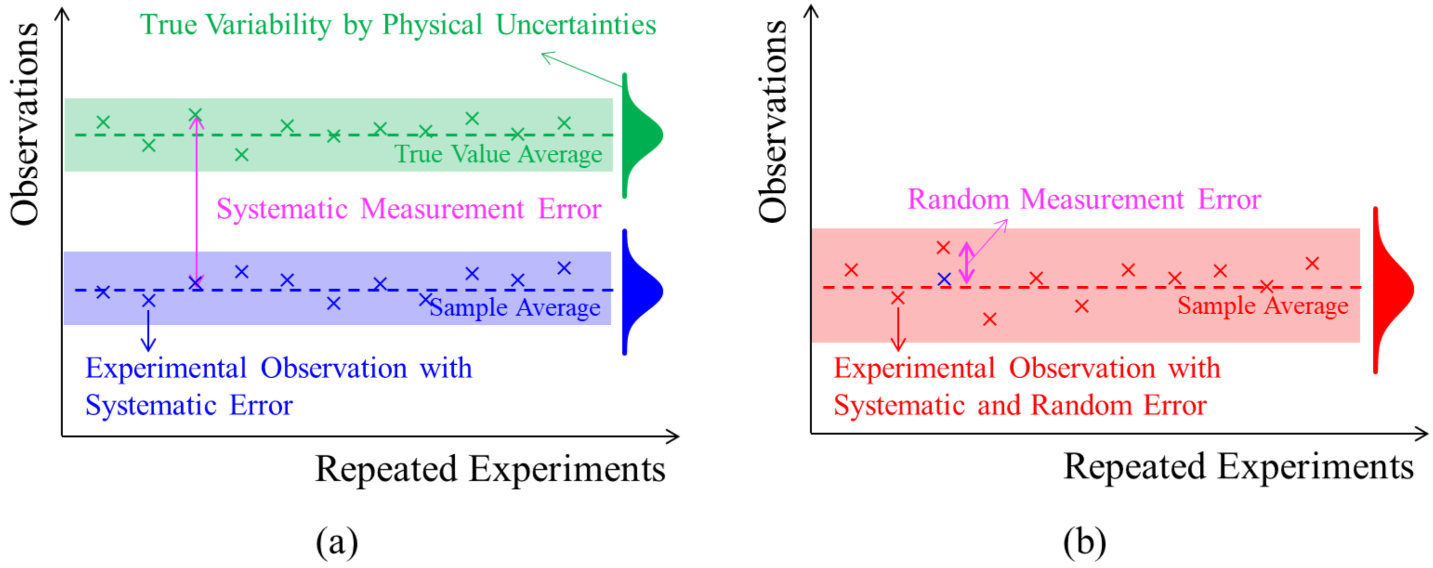

Figure 2-1 Illustration of Measurement Errors: Observed Data (a) With Systematic Measurement Error and (b) With Systematic and Random Measurement Errors

2.1.3 Characterization of Measurement Errors

The major difference between the two types of measurement errors is that a particular source of a systematic measurement error will affect all the replicate measurements in the same manner. In contrast, a source of random measurement error will randomly affect the measurements. Generally, for this reason, the uncertainties that arise due to systematic and random measurement errors are characterized by a uniform distribution and a normal distribution, respectively. This section presents how to characterize the two measurement errors using probability distributions.

Figure 2-2 describes probability distributions for systematic and random measurement errors for given observed data (y). First, systematic measurement errors, which bias the measurement from the true value, emerge from a source of measurement error that is effectively the same in all the replicate tests. This is described by a uniform distribution that is defined by a lower bound and an upper bound (Figure 2-2a). One of the two bounds is set to '0', which means there is no bias from the true value. The other bound denotes the maximum amount of bias 'α', relative to the measured value (y). For a negative systematic error (Figure 2-2a, where a negative systematic error biases the true value to a smaller measured value), the upper bound is set to '0' and the lower bound is set to '-α' relative to the measured value (y), and vice-versa for a positive systematic error ([0, αy]). Second, the random measurement errors are characterized to follow normal distributions due to the indefinite multiple sources of the randomness. The random measurement error, which follows a normal distribution with '0' mean and 'βy' standard deviation, is added to the value biased by the systematic measurement error. Note that the random measurement error can cause either an increase or a decrease in the estimate of the true value.

Figure 2-3 describes the PDFs of the true value with measurement errors ($y + e$). First, the gray-shaded PDFs ($p(y)$) in the upper part of Figure 2-3 denote the PDF of the true value (y) without any measurement error. The proposed method aims to obtain $p(y)$, which describes the uncertainty in the observed data without measurement errors. Suppose that y_1 and y_2 are arbitrarily sampled from $p(y)$. Depending on the randomly sampled values (y_1 and y_2), the PDFs of e are defined using the method explained in Figure 2-3 ($p(e_{sys}|y)$, $p(e_{ran}|y)$). Then, the PDFs of the true value with two measurement errors ($y + e_{sys}$, $y + e_{ran}$) are formulated as the bottom parts of Figure 2-3a and 2-3b, respectively. In Figure 2-3a, two uniform distributions are defined by the lower bounds ($y_1 - \alpha y_1$, $y_2 - \alpha y_2$) and upper bounds (y_1, y_2) for two samples, y_1 and y_2. This explains that an observation y_1 with systematic measurement error (e_{sys}) can be observed from the left blue-shaded uniform distribution. Therefore, an uncertainty characterization process is required to eliminate the biased effect from the systematic measurement error (e_{sys}) to obtain the gray-shaded true population $p(y)$. In Figure 2-3b, two normal distributions are defined by the means (y_1, y_2) and standard deviations ($\beta_1 y_1$, $\beta_2 y_2$) of the two samples, y_1 and y_2. An observation y_1 with a random measurement error (e_{ran}) can be observed from the left red-shaded uniform distribution. Therefore, an uncertainty characterization process is required to eliminate the random effect that arises from the random measurement error (e_{ran}) to obtain the gray-shaded true population $p(y)$. For a case where both systematic and random measurement errors exist, $p(y + e|y)$ can be defined by integrating the systematic (e_{sys}) and random measurement (e_{ran}) errors into e. Based on the characterization of uncertainties and measurement errors in this section, the following section introduces the step-by-step mathematical formulations for the proposed method.

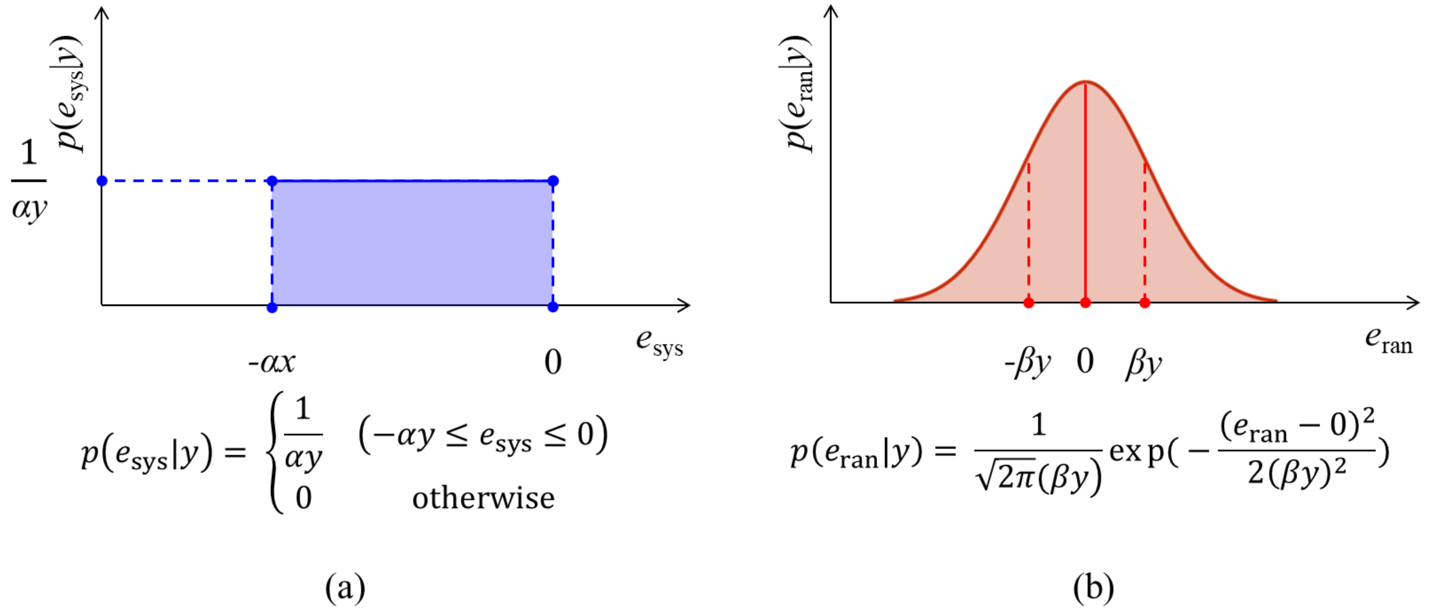

Figure 2-2 Probability Density Function of Measurement Errors: (a) Systematic Measurement Error, (b) Random Measurement Error

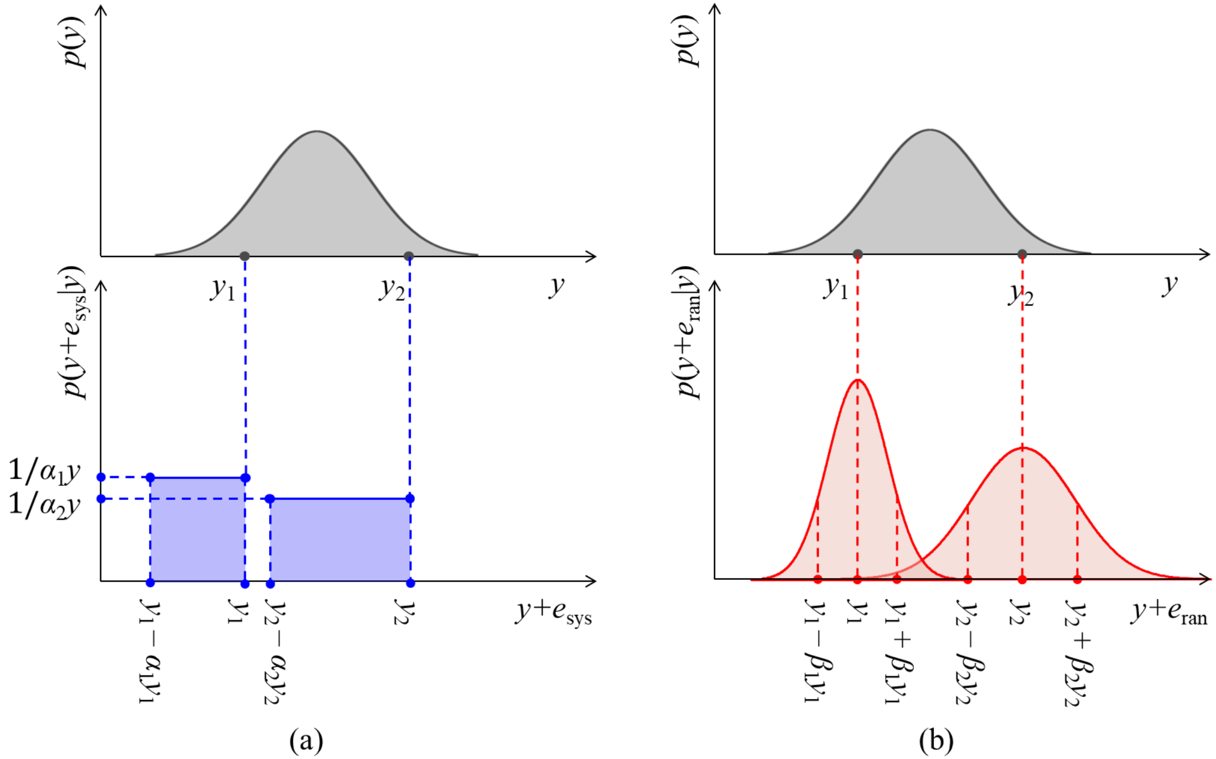

Figure 2-3 Probability Density Function (PDF) of the True Value (y) of the Quantity of Interest, with Measurement Errors (e): (a) PDF of y and Systematic Measurement Error (e_{sys}), (b) PDF of y and Random Measurement Error (e_{ran})

2.2 Uncertainty Characterization of Experimental Observations Using the Maximum Likelihood Estimation

The objective of Research Thrust 1 is to estimate the statistical parameters of an assumed parametric probability distribution ($p(y)$) that describes the variability and uncertainties of the QOI using n numbers of experimental observations ($\mathbf{Y} = \{y_1, \ldots, y_n\}$). The proposed method uses maximum likelihood estimation to estimate the values of the statistical parameters that maximize the formulated likelihood function. This section outlines the overall steps for characterizing the uncertainty of experimental observations using maximum likelihood estimation. Section 2.2.1 develops the probability density function (PDF) for the variability in observed system responses, which results from the physical uncertainties and measurement errors. Section 2.2.2 formulates the likelihood function using developed and assumed PDFs for the observed system response and measurement errors. Section 2.2.3 derives the mathematical formulation for obtaining the values of the statistical parameters that maximize the formulated likelihood function.

2.2.1 Step 1: Development of Parametric Probability Distributions for Describing the Uncertainty in Experimental Observations Considering Measurement Errors

First, Equation (2-1) explains how the observed data (y_{obs}) includes the true variability (y_{true}) and measurement errors (e), which are composed of systematic (e_{sys}) and random (e_{ran}) measurement errors.

$$y_{\mathrm{obs}} = y_{\mathrm{true}} + e = y_{\mathrm{true}} + e_{\mathrm{sys}} + e_{\mathrm{ran}} \tag{2-1}$$

To describe the true variability of experimental observations (y_{true}), a Gaussian distribution is assumed with statistical parameters ((the mean ($\mu_{y\mathrm{true}}$) and standard deviation ($\sigma_{y\mathrm{true}}$)) to be estimated later (Equation (2-2)).

$$\begin{aligned} p(y_{\mathrm{true}}) &= N(\mu_{y_{\mathrm{true}}}, \sigma_{y_{\mathrm{true}}}) \\ &= \frac{1}{\sqrt{2\pi}\sigma_{y_{\mathrm{true}}}} \exp\left(-\frac{\left(y_{\mathrm{true}} - \mu_{y_{\mathrm{true}}}\right)^2}{2\sigma_{y_{\mathrm{true}}}{}^2}\right) \end{aligned} \tag{2-2}$$

Based on the information given in previous section, a uniform and normal distribution are formulated for the probability distributions of the systematic and random measurement errors for a given system response (y), respectively (Equations (2-3) and (2-4).

$$p(e_{\mathrm{sys}}|y) = \begin{cases} \dfrac{1}{\alpha y} & (-\alpha y \le e_{\mathrm{sys}} \le 0) \\ 0 & \text{otherwise} \end{cases} \tag{2-3}$$

$$p(e_{\mathrm{ran}}|y) = N(0, \beta y) = \frac{1}{\sqrt{2\pi}(\beta y)} \exp\left(-\frac{(e_{\mathrm{ran}} - 0)^2}{2(\beta y)^2}\right) \tag{2-4}$$

Then, by defining the integrated error (e) as $e = e_{\mathrm{sys}} + e_{\mathrm{ran}}$, the probability distribution of e for a given y can be presented as the convolution of the systematic and random measurement errors, as shown in Equations (2-5) and (2-6).

$$p(e|y) = \int p(e - e_{\mathrm{sys}}|y)\, p(e_{\mathrm{sys}}|y)\, de_{\mathrm{sys}} \tag{2-5}$$

$$p(e|y) = \frac{1}{\alpha y} \int_{-\alpha y}^{0} \frac{1}{\sqrt{2\pi}(\beta y)} \times \exp\left(-\frac{\left(d - y - e_{\mathrm{sys}}\right)^2}{2(\beta y)^2}\right) de_{\mathrm{sys}} \tag{2-6}$$

Using Equation (2-1), the marginal probability distribution of the observed system response (y_{obs}) for a given y can be formulated as

$$p_{y_{\mathrm{obs}}|y}(y_{\mathrm{obs}}|y) = p_{e|y}(y_{\mathrm{obs}} - y|y)$$
$$= \frac{1}{\alpha y}\int_{-\alpha y}^{0} N(e_{\mathrm{sys}}|y_{\mathrm{obs}} - y, \beta y)de_{\mathrm{sys}} \quad (2\text{-}7)$$

Then, using the sum and product rule, the marginal distribution ($p(y)$) can be expressed as

$$p(y_{\mathrm{obs}}) = \int p(y_{\mathrm{obs}}|y)p(y)dy \quad (2\text{-}8)$$

The following section introduces how to develop the likelihood function based on the results of this section.

2.2.2 Step 2: Formulation of the Likelihood Function

In general, the most effective way to characterize the uncertainties in an engineered system is to represent the uncertainty in the form of a PDF. Parametric methods are common for a PDF to be parameterized or characterized by statistical parameters, for example, the mean and standard deviation of a Gaussian distribution (Yates 1934) (Smirnov 1948) (Massey Jr 1951) (Anderson and Darling 1952) (Anderson and Darling 1954) (Stephens 1974) (Plackett 1983). As long as a parametric PDF for a system response can be clearly specified, parametric methods are without doubt powerful tools for characterizing uncertainties. A parametric method requires two steps: 1) selecting a proper type of PDF to describe the uncertainty, and; 2) estimating values for statistical parameters of the selected PDF. In this study, the first step is completed assuming that the type of probability distribution for the system responses follows a Gaussian distribution. After an appropriate PDF is determined, statistical parameters can be estimated by several methods, including maximum likelihood estimation (Charnes et al. 1976) (Scholz 1985) (Newey and West 1987) (Myung 2003).

Maximum likelihood estimation can be used to estimate the unknown statistical parameters of a parametric probability distribution (McLachlan and Krishnan 2007). Maximum likelihood estimation searches for proper values of the statistical parameters that maximize the likelihood function for given the observations. This implies that using maximum likelihood estimation requires building of the likelihood function. Therefore, this section presents how to construct the likelihood function. Equation (2-9) denotes the likelihood function (L) for a given n number of observations ($\mathbf{Y} = \{y_1, \ldots, y_n\}$).

$$L(y) = \prod_{i=1}^{n} p(y_i) = \prod_{n} \int p(y_n|y)p(y)dy \tag{2-9}$$

Since the type of probability distribution for a system response is assumed to follow a Gaussian distribution in this study, Equation (2-9) can be reformulated as

$$L(y|\mu,\sigma) = \prod_{n} \int p(y_n|y)N(\mu_y,\sigma_y)dy \tag{2-10}$$

Using the logarithm, the log-likelihood function can then be formulated as

$$\ln(L(D|\mu,\sigma)) = \sum_{n} \ln(\int p(y_n|y)N(\mu_y,\sigma_y)dy) \tag{2-11}$$

2.2.3 Step 3: Maximization of the Likelihood Function

Based on the work described in Section 2.2.2, this section computes derivatives of Equation (2-11) with respect to the mean (μ_y) and standard deviation (σ_y). By equating the derivative of Equation (2-11) to zero, an analytical solution for the value of the mean (μ_y) and standard deviation (σ_y) can be obtained. The derived results are shown as Equations (2-12) and (2-13):

$$\mu_y = \frac{1}{n}\sum_n \frac{\int_{-\infty}^{\infty} y p(y_n|y) N(\mu_y, \sigma_y) dy}{\int_{-\infty}^{\infty} p(d_n|y) N(\mu_y, \sigma_y) dy} \tag{2-12}$$

$$\sigma_y^2 = \frac{1}{n}\sum_n \frac{\int_{-\infty}^{\infty} (y - \mu_y)^2 p(y_n|y) N(\mu_y, \sigma_y) dy}{\int_{-\infty}^{\infty} p(y_n|y) N(\mu_y, \sigma_y) dy} \tag{2-13}$$

Finally, by substituting Equations (2-2) and (2-7) into N(μ_y, σ_y) and $p(y_n|y)$, respectively, in Equations (2-12) and (2-13), the analytical estimate of μ_y and σ_y^2 can be obtained. However, it should be noted that Equations (2-12) and (2-13) are not explicit; this means that a final estimate is required to calculate the estimate. In this case, the estimate can be obtained by updating iteratively until convergence.

2.3 Summary and Discussion

With an aim to characterize the uncertainty in experimental observations for a system response of an engineered system, Research Thrust 1 utilized maximum likelihood estimation to estimate the unknown statistical parameters of an assumed probability distribution. First, the likelihood function was formulated with characterization of the variability by physical uncertainties, systematic measurement error, and random measurement error. To prove the effectiveness of the proposed idea, a case study "cantilever beam problem" (introduced in Section 6.1: Case Study 1) is used to show that the proposed method produces accurate estimation when measurement errors exist in the observations. Also, by adopting the method to the case studies "an automobile steering column problem (Section 6.3: Case Study 3)" and "a liquid crystal display panel problem (Section 6.4: Case Study 4)," the proposed method is shown to be applicable to a real-world engineering problem. In Case Study 3, an automobile steering column problem, the proposed method is adopted to consider the systematic measurement error due to different experimental conditions and the random measurement error due to human errors for accurate characterization of the true variability in observations. In Case Study 4, a LCD panel problem, the proposed method is used to gather the observation data from different experimental conditions. With the proposed idea, statistical uncertainty by lack of data could be eased by increased numbers of experimental observations for the same experimental condition.

To expand use of the proposed method, it will be challenging to determine the value of statistical parameters for the probability distribution of systematic and random measurement errors. For example, in this study, the maximum amount of bias 'α' of systematic measurement error and 'βy' standard deviation for random measurement are considered to be given information. However, information about those parameters may not be available in real-world applications. Thus, there is a need for a systematic way to determine the amount of systematic and random measurement errors present in real-world settings.

In addition, this chapter introduces only the Gaussian case. However, when using a parametric approach, it should be noted that an incorrect assumption about the type of probability distribution may cause numerical error, or statistical uncertainty. In case the type of probability distribution is known to be non-Gaussian, the mathematical derivations from Equation (2-2) to Equation (2-13) should be re-derived. However, the mathematical derivation process can be extended to non-Gaussian types of probability distributions.

Sections of this chapter have been published or submitted as the following journal article:

1) Taejin Kim, **Guesuk Lee (co-first author)**, and Byeng D. Youn, "Uncertainty Characterization Under Measurement Errors Using Maximum Likelihood Estimation: Cantilever Beam End-to-End UQ Test Problem," *Structural and Multidisciplinary Optimization*, Accepted in November 2018.

Chapter 3

Robust Optimization-Based Model Calibration with Analytical Sensitivity Information

Optimization-based model calibration (OBMC) adopts optimization techniques to estimate unknown input variables of a computational model in a probabilistic sense. The fields of design optimization have developed various optimization algorithms; these can be broadly classified as gradient-based (local) and global (non-local or evolutionary) algorithms (Arora 2004) (Tekin and Sabuncuoglu 2004) (Snyman 2005). Compared to global algorithms, gradient-based algorithms require a relatively small number of function evaluations to find the optimum point (calibrated point). Therefore, prior studies on OBMC have primarily used gradient-based algorithms due to the efficiency issue (Youn et al. 2011) (Fender et al. 2014) (Jung et al. 2014) (Jung et al. 2016).

A calibration metric, which is defined as an objective function of OBMC, is an important element of OBMC. Various types of calibration metrics have been developed to quantify the difference or the similarity between the system responses derived from computational predictions and experimental observations, such as the likelihood function, the probability residual, and others (Oberkampf et al. 2004a) (McFarland and Mahadevan 2008a) (McFarland and Mahadevan 2008b) (Youn et al. 2011) (Jung et al. 2014) (Choi et al. 2016) (Oh et al. 2017). Based on the author's experience, however, OBMC that uses existing calibration metrics has been shown to lead to instable calibration processes and inaccurate calibrated results (e.g., divergence or convergence at large standard deviations or biased mean values, as depicted in Figure 1-5).

There are four possible reasons that cause unstable and inaccurate OBMC when gradient optimization algorithms and existing calibration metrics are used, including: 1) non-convexity of the optimization problem when OBMC is performed with existing calibration metrics; 2) existence of a local minimum; 3) an extremely low degree of sensitivity, and; 4) randomness due to sampling methods used for the uncertainty propagation process (Figure 3-1). Thus, Research Thrust 2 first investigates the global convexity of OBMC optimization problems with calibration metrics. In order to check the global convexity, the first- and second-order derivatives (sensitivity information) of the objective functions (calibration metrics) with respect to the calibration parameters are derived. This analytical derivation includes several assumptions. Next, a condition for convex optimization (e.g., a determinant of the Hessian matrix) is investigated using the derivatives. Using the analytical sensitivity information, other remaining possible causes are investigated. Finally, OBMC with derived analytical sensitivity information is performed, showing a robust optimization process.

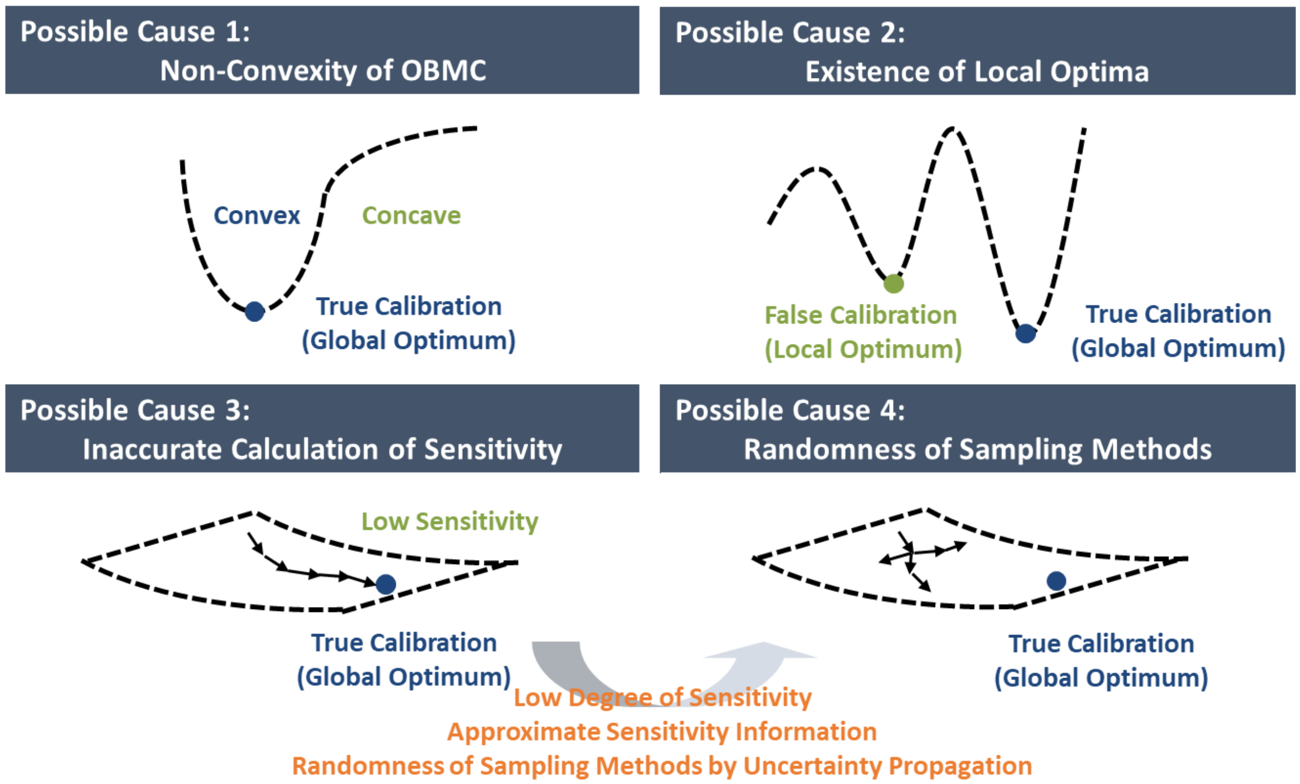

Figure 3-1 Possible Causes of Inaccurate and Unstable Calibrated Results Derived Optimization-Based Model Calibration with Existing Calibration Metrics

The remainder of Chapter 3 is organized as follows. Section 3.1 summarizes mathematical formulations used to check the global convexity for the optimization problem. Section 3.2 reviews two existing calibration metrics: 1) the likelihood function and 2) the probability residual. In Section 3.3, a comprehensive analysis of OBMC using introduced calibration metrics is given; both analytical results and numerical example results are provided. Later, in Section 6.2, a case study of a "bearing capacity equation" is introduced to check the effectiveness of the proposed idea; this case study executes OBMC using the analytical sensitivity information derived in previous sections. Section 3.4 presents an overview of robust OBMC with analytical sensitivity information. Finally, the conclusions of this work are provided in Section 3.5.

3.1 Conditions of a Convex Optimization for Optimization-Based Model Calibration

For the optimization algorithms used in OBMC, gradient-based algorithms (e.g., the gradient descent method, Newton's methods, or sequential quadratic programming) are recommended due to their efficiency. However, gradient methods do not ensure that the subsequence of iterations converges to the global optimum unless it is confirmed to be a convex optimization problem. If OBMC with calibration metrics is confirmed to be convex, it can be solved efficiently and accurately. Furthermore, a convex optimization converges to the global optimum regardless of the location of the initial point. In summary, recognizing or formulating a convex optimization is a great advantage when solving an optimization problem, including OBMC. Thus, this section provides and overview of a way to check the convexity of an optimization problem. First, Section 3.1.1 provides a general way to check global convexity of a function. Section 3.1.2 presents how to derive analytical sensitivity information to check the global convexity under assumptions.

3.1.1 Analysis of Global Convexity for a Function

For a function of several variables, $f(\mathbf{x})$, where $\mathbf{x}$ is an n-vector, the multidimensional Taylor's expansion at the given point $\mathbf{x}^*$ can be formulated as

$$f(\mathbf{x}) = f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d} + \frac{1}{2}\mathbf{d}^T\mathbf{H}(\mathbf{x}^*)\mathbf{d} + R \tag{3-1}$$

where $\mathbf{x} - \mathbf{x}^* = \mathbf{d}$, $\mathbf{H}$ is the n×n Hessian matrix, and R is the remainder term. Alternatively, a change in the function $\Delta f = f(\mathbf{x}) - f(\mathbf{x}^*)$ is given as

$$\Delta f = \nabla f(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d} + \frac{1}{2}\mathbf{d}^T\mathbf{H}(\mathbf{x}^*)\mathbf{d} + R \tag{3-2}$$

If $\mathbf{x}^*$ is a local minimum for a convex function, then Δf must be non-negative. Δf can be non-negative for all possible $\mathbf{d}$ unless $\nabla f(\mathbf{x}^*) = 0$. Considering the second term, the positivity of Δf is assured if $\mathbf{d}^{\mathrm{T}}\mathbf{H}(\mathbf{x}^*)\mathbf{d} > 0$. This is true if the Hessian $\mathbf{H}(\mathbf{x}^*)$ is a positive definite matrix. For a twice differentiable f, the positive definite of the $\mathbf{H}(\mathbf{x}^*)$ can be confirmed by

$$\det \mathbf{H}\big(f(\mathbf{x}^*)\big) > 0 \text{ or } \geq 0 \tag{3-3}$$

where > 0 denotes positive definite and ≥ 0 denotes positive semidefinite. For a two-variable case, a function is convex if and only if

$$\frac{\partial^2 f}{\partial \mathrm{x}_1^2}(x)\frac{\partial^2 f}{\partial \mathrm{x}_2^2}(x) - \frac{\partial^2 f}{\partial x_1 \partial x_2}(x)\frac{\partial^2 f}{\partial x_2 \partial x_1}(x) > 0 \text{ or } \geq 0 \tag{3-4}$$

$$\text{and} \quad \frac{\partial^2 f}{\partial x_1^2}(x), \frac{\partial^2 f}{\partial x_2^2}(x) > 0 \text{ or } \geq 0 \quad \text{for} \quad all \ \ x \ \ \in A$$

where > 0 denotes positive definite and ≥ 0 denotes positive semidefinite. If the above (the Hessian matrix of the function (*f*) is positive semidefinite or positive definite) is checked for all points in set *A*, then the function can be called globally convex (strictly convex for positive definite) in set *A*.

3.1.2 Derivation of Analytical Sensitivity Information for OBMC Under Assumptions

Following the work described in Section 3.1.1, it is required to obtain the first-order derivatives (gradient vector) and the Hessian matrix composed of the second-order derivatives to check the convexity of OBMC using calibration metrics. For a continuous and at least twice continuously differentiable function (*f*, the calibration metric in this study), the first- and second-order derivatives can be obtained by using the chain rule and the product rule, as follows

$$\frac{\partial f}{\partial \boldsymbol{\theta}_{\mathbf{X}_{\text{unknown}}}} = \frac{\partial f}{\partial \boldsymbol{\theta}_y} \times \frac{\partial \boldsymbol{\theta}_y}{\partial \boldsymbol{\theta}_{\mathbf{X}_{\text{unknown}}}} \tag{3-5}$$

$$\frac{\partial^2 f}{\partial \boldsymbol{\theta}^2_{\mathbf{X}_{\text{unknown}}}} = \frac{\partial^2 f}{\partial \boldsymbol{\theta}_y \, \partial \boldsymbol{\theta}_{\mathbf{X}_{\text{unknown}}}} \times \frac{\partial \boldsymbol{\theta}_y}{\partial \boldsymbol{\theta}_{\mathbf{X}_{\text{unknown}}}} + \frac{\partial f}{\partial \boldsymbol{\theta}_y} \times \frac{\partial^2 \boldsymbol{\theta}_y}{\partial \boldsymbol{\theta}^2_{\mathbf{X}_{\text{unknown}}}} \tag{3-6}$$

By assuming that, 1) there are multiple known input variables, one unknown input variable, and one system response (output), 2) the inputs ($\mathbf{X}$ = {$\mathbf{X}_{\text{known}}$, x_{unknown}}) follow normal distribution, 3) the relationship (*M*, model) between the inputs (**X**) and one output (*y*) is linear, and 4) the output (*y*) follows normal distribution, Equations (3-5) and (3-6) can be reformulated with two statistical parameters (i.e., mean ($\mu_{x_{\text{unknown}}}$) and standard deviation ($\sigma_{x_{\text{unknown}}}$)) of unknown input variables as follows

$$\frac{\partial f}{\partial \mu_{x_{\text{unknown}}}} = \frac{\partial f}{\partial \mu_y} \times \frac{\partial \mu_y}{\partial \mu_{x_{\text{unknown}}}} \tag{3-7}$$

$$\frac{\partial f}{\partial \sigma_{x_{\text{unknown}}}} = \frac{\partial f}{\partial \sigma_y} \times \frac{\partial \sigma_y}{\partial \sigma_{x_{\text{unknown}}}} \tag{3-8}$$

$$\frac{\partial^2 f}{\partial \mu^2_{x_{\text{unknown}}}} = \frac{\partial^2 f}{\partial \mu^2_y} \times \frac{\partial \mu_y}{\partial \mu_{x_{\text{unknown}}}} \times \frac{\partial \mu_y}{\partial \mu_{x_{\text{unknown}}}} + \frac{\partial f}{\partial \mu_y} \times \frac{\partial^2 \mu_y}{\partial \mu^2_{x_{\text{unknown}}}} \tag{3-9}$$

$$\frac{\partial^2 f}{\partial \sigma^2_{x_{\text{unknown}}}} = \frac{\partial^2 f}{\partial \sigma^2_y} \times \frac{\partial \sigma_y}{\partial \sigma_{x_{\text{unknown}}}} \times \frac{\partial \sigma_y}{\partial \sigma_{x_{\text{unknown}}}} + \frac{\partial f}{\partial \sigma_y} \times \frac{\partial^2 \sigma_y}{\partial \sigma^2_{x_{\text{unknown}}}} \tag{3-10}$$

$$\frac{\partial^2 f}{\partial \mu_{x_{\text{u}}} \partial \sigma_{x_{\text{u}}}} = \frac{\partial^2 f}{\partial \mu_y \partial \sigma_y} \times \frac{\partial \sigma_y}{\partial \sigma_{x_{\text{u}}}} \times \frac{\partial \mu_y}{\partial \mu_{x_{\text{u}}}} + \frac{\partial f}{\partial \mu_y} \times \frac{\partial^2 \mu_y}{\partial \mu^2_{x_{\text{u}}}} \tag{3-11}$$

Since the relationship between the inputs (**X**) and the output (y) is linear, for example,

$$y = a_1 x_1 + a_2 x_2 + \cdots + a_n x_N + b \tag{3-12}$$

$$\frac{\partial f}{\partial \sigma_{x_{\text{unknown}}}} = \frac{\partial f}{\partial \sigma_y} \times \frac{\partial \sigma_y}{\partial \sigma_{x_{\text{unknown}}}} \tag{3-13}$$

$$\mu_y = a_1 \mu_{x_1} + a_2 \mu_{x_2} + \cdots + a_N \mu_{x_N} + b \tag{3-14}$$

$$\frac{\partial \mu_y}{\partial \mu_{x_{\text{u}}}} = a_{\text{u}} \;\text{ and }\; \frac{\partial^2 \mu_y}{\partial \mu^2_{x_{\text{u}}}} = 0 \tag{3-15}$$

$$\sigma^2_y = {a_{\text{u}}}^2 \sigma^2_x \;\; (\sigma_y = |a_{\text{u}}| \sigma_x) \tag{3-16}$$

where N is the total number of input variables, including known and unknown input variables, and u denotes the index of the unknown input variable that needs to be calibrated. a_i (i = 1, …, N) and b are respectively the coefficients of inputs and the constant term of the linear problem. a_{u}, which is one of the coefficients of inputs, denotes the linear coefficient for the unknown input variable. Thus, equations from (3-7) to (3-11) can be reformulated as

$$\frac{\partial f}{\partial \mu_{x_\mathrm{u}}} = \frac{\partial f}{\partial \mu_y} \times a_\mathrm{u} \tag{3-17}$$

$$\frac{\partial f}{\partial \sigma_{x_\mathrm{u}}} = \frac{\partial f}{\partial \sigma_y} \times |a_\mathrm{u}| \tag{3-18}$$

$$\frac{\partial^2 f}{\partial \mu_{x_\mathrm{u}}^2} = \frac{\partial^2 f}{\partial \mu_y^2} \times {a_\mathrm{u}}^2 \tag{3-19}$$

$$\frac{\partial^2 f}{\partial \sigma_{x_\mathrm{u}}^2} = \frac{\partial^2 f}{\partial \sigma_y^2} \times {a_\mathrm{u}}^2 \tag{3-20}$$

$$\frac{\partial^2 f}{\partial \mu_{x_\mathrm{u}} \partial \sigma_{x_\mathrm{u}}} = \frac{\partial^2 f}{\partial \mu_y \partial \sigma_y} \times |a_\mathrm{u}| \times a_\mathrm{u} \tag{3-21}$$

The next section introduces existing calibration metrics with mathematical formulations. Later, using the equation derived in this section, the convexity of OBMC with those calibration metrics is investigated.

3.2 Brief Review of Calibration Metrics for OBMC

Prior research efforts have developed calibration metrics for calibrating unknown input variables. Among them, two popular calibration metrics are introduced in this section: 1) the likelihood function (Section 3.2.1) and 2) the probability residual (Section 3.2.2).

3.2.1 The Likelihood Function

The likelihood function is one frequently used method for estimating statistical parameters in statistics (White 1982) (Myung 2003). The likelihood function has also been used for OBMC, which aims at estimating the statistical parameters of unknown input variables (Youn et al. 2011) (Fender et al. 2014) (Jung et al. 2014) (Jung et al. 2016). The likelihood function for OBMC is defined as

$$L(\mathbf{Y}_{\text{obs}}|\boldsymbol{\theta}_{\mathbf{X}}) = L\left(y_1, y_2, \dots, y_n \middle| \boldsymbol{\theta}_{\mathbf{X}_\text{k}}, \boldsymbol{\theta}_{\mathbf{X}_\text{u}}\right) = \prod_{i=1}^{n} p_{\mathbf{Y}_{\text{pre}}}(y_i|\boldsymbol{\theta}_{\mathbf{x}}) \tag{3-22}$$

where $p_{\mathbf{Y}_{\text{pre}}}(\cdot)$ denotes the PDF of predicted system responses described by a vector of statistical parameters ($\boldsymbol{\theta}_{\mathbf{X}}$) of both known and unknown input variables. y_i denotes the i-th observed system response from a set of n measurements in experiments. For given statistical parameters ($\boldsymbol{\theta}_{\mathbf{X}}$), the likelihood ($L(\mathbf{Y}_{\text{obs}}|\boldsymbol{\theta}_{\mathbf{X}})$) is evaluated using a set of independent experimental observations (y_i). The likelihood value shows a maximum value when the predicted PDF ($p_{\mathbf{Y}_{\text{pre}}}(\cdot)$) with a given set of the statistical parameters ($\boldsymbol{\theta}_{\mathbf{X}}$) shows utmost agreement with the experimental data (y_i). For OBMC, the statistical parameters ($\boldsymbol{\theta}_{\mathbf{X}}$) of unknown input variables are calibrated by maximizing the likelihood function value. Frequently, in practical use, the natural logarithm of the likelihood function is convenient to work with; it can be reformulated as

$$\begin{aligned} L(\mathbf{Y}_{\text{obs}}|\boldsymbol{\theta}_{\mathbf{X}}) &= L\left(y_1, y_2, \dots, y_n \middle| \boldsymbol{\theta}_{\mathbf{X}_\text{k}}, \boldsymbol{\theta}_{\mathbf{X}_\text{u}}\right) \\ &= \sum_{i=1}^{n} \ln(p_{\mathbf{Y}_{\text{pre}}}(y_i|\boldsymbol{\theta}_{\mathbf{x}})) \end{aligned} \tag{3-23}$$

Since the logarithm is a monotonically increasing function, the result estimated by the log-likelihood function is the same as the one determined by the likelihood function itself. By taking the natural logarithm, a product of individual likelihood functions is converted to a sum of individual logarithms. This allows convenient computation of the derivations. In the next section, the analytical sensitivity information of the likelihood function is examined, which involves taking the first-order and second-order derivatives of the likelihood function. Therefore, the log-likelihood formulation is used for this study. One drawback of the likelihood function is that it can produce a negative infinite number when there is a big difference between the experimental data and the predicted PDF. Mathematically, the value of ($p_{\mathbf{Y}\mathrm{pre}}(y_i|\boldsymbol{\theta}_\mathbf{X})$) can have an extremely small value for the predicted PDF described by $\boldsymbol{\theta}_\mathbf{X}$, of which the mean value is too biased from the expected optimum point, and the standard deviation is too small or large.

3.2.2 The Probability Residual

To quantify the differences between distributions of the simulation and experimental data, the probability residual (PR) (Choi et al. 2016) (Oh et al. 2017) is formulated either by the integral of the absolute difference or the squared Euclidean distance, which is defined as

$$PR_{\mathrm{abs}}(\mathbf{Y}_{\mathrm{obs}}|\boldsymbol{\theta}_\mathbf{X}) = \int_{-\infty}^{\infty} |p_{\mathbf{Y}_{\mathrm{obs}}}(y) - p_{\mathbf{Y}_{\mathrm{pre}}}(y|\boldsymbol{\theta}_\mathbf{X})| dy \tag{3-24}$$

$$PR_{\mathrm{squ}}(\mathbf{Y}_{\mathrm{obs}}|\boldsymbol{\theta}_\mathbf{X}) = \int_{-\infty}^{\infty} \left(p_{\mathbf{Y}_{\mathrm{obs}}}(y) - p_{\mathbf{Y}_{\mathrm{pre}}}(y|\boldsymbol{\theta}_\mathbf{X})\right)^2 dy \tag{3-25}$$

where $p_{\mathbf{Y}\mathrm{obs}}(y)$ denotes the PDF of the system response from experimental observation and $p_{\mathbf{Y}\mathrm{pre}}(y|\boldsymbol{\theta})$ denotes the PDF of the system response from computational prediction. For both types of the PR, the minimum value 'zero' of the PR is obtained when the predicted PDF from a given set of the statistical parameters shows the best agreement with the observed PDF. Compared with the likelihood function, the PR requires an uncertainty characterization process for the PDF ($p_{\mathbf{Y}\mathrm{obs}}(y)$) of experimental observation, which may cause statistical uncertainty, since a dearth of observed data can lead to erroneous characterization of the PDF. One merit of the PR is that accurate one-to-one comparison between two probability distributions from observations and predictions is available. For accurate OBMC, this is a desired characteristic.

It is noteworthy that the formulation of the PR is based on a quadratic function, which is normally convex for the entire domain. However, as examined in the next section, OBMC with PR is not strictly convex for the entire domain of optimization variables (calibration parameters).

3.3 Comprehensive Investigation of OBMC with Calibration Metrics

The first objective of Research Thrust 2 is to examine whether OBMC with existing calibration metrics formulates a convex optimization problem. This section presents analytical (Section 3.3.1) and numerical (Section 3.3.2) investigations for checking the convexity of OBMC, using the two previously introduced calibration metrics.

3.3.1 Analytical Investigation

Analytical investigation of OBMC with two calibration metrics is conducted with four assumptions: 1) the model is in a linear relationship, 2) only one unknown input variable exists in the model, 3) the uncorrelated input variables of the model follow a normal distribution, and 4) the simulation and experimental data both follow normal distributions.

To check the convexity of OBMC with the likelihood function (Section 3.3.1.1) and PR (Section 3.3.1.2) (in this study, the probability residual based on the integral of squared difference is adopted), the first-order derivatives (gradient vector) and the

Hessian matrix composed of the second-order derivatives are derived. A major proposition is the assumption that the PDFs ($p_{\mathbf{Y}\text{pre}}(y|\boldsymbol{\theta})$, $p_{\mathbf{Y}\text{obs}}(y)$) in Equations (3-23) and (3-25) follow the normal distribution described with the mean and standard deviation (or variance). Another important assumption is that a linear relationship exists between inputs and an output.

3.3.1.1. Analytical Sensitivity Information of the Likelihood Function

First, OBMC using the likelihood function (L) is a maximization problem. In this study, for consistency with the other metric, the maximization problem is reversed to minimization by multiplying by '-1'. With the aforementioned assumptions, the likelihood function in Equation (3-23) can be reformulated as

$$L(\mathbf{Y}_{\text{obs}}|\boldsymbol{\theta}_{\mathbf{X}}) = \frac{1}{2{\sigma_y}^2}\sum_{i=1}^{n}\left(y_i - \mu_y\right)^2 + \frac{n}{2}\ln{\sigma_y}^2 + \frac{n}{2}\ln 2\pi \tag{3-26}$$

Let variance $s_y = {\sigma_y}^2$ ($s_y > 0$). The first derivative (gradient vector) of L with respect to two statistical parameters, the mean (μ_y) and variance (s_y) of the predicted PDF, is represented as follows

$$\begin{bmatrix} \dfrac{\partial L}{\partial \mu_y} \\ \dfrac{\partial L}{\partial s_y} \end{bmatrix} = \begin{bmatrix} -\dfrac{1}{s}\sum_{i=1}^{n}(y_i - \mu_y) \\ -\dfrac{1}{2s^2}\sum_{i=1}^{n}\left(y_i - \mu_y\right)^2 + \dfrac{N}{2s} \end{bmatrix} \tag{3-27}$$

Differentiating the gradient vector once again, a matrix of the second partial derivatives for the likelihood function with respect to two statistical parameters – mean (μ_y) and variance (s_y) of the predicted PDF – can be obtained as

$$
\begin{pmatrix} \frac{\partial^2 f}{\partial \mu_y^2} & \frac{\partial^2 L}{\partial \mu_y \partial s_y} \\ \frac{\partial^2 L}{\partial s_y \partial \mu_y} & \frac{\partial^2 f}{\partial s_y^2} \end{pmatrix}
= \begin{pmatrix} \frac{n}{s_y} & \frac{1}{{s_y}^2}\sum_{i=1}^{n}(y_i - \mu_y) \\ \frac{1}{{s_y}^2}\sum_{i=1}^{n}(y_i - \mu_y) & \frac{1}{{s_y}^3}\sum_{i=1}^{n}\left(y_i - \mu_y\right)^2 - \frac{n}{2{s_y}^2} \end{pmatrix} \tag{3-28}
$$

Using the Equations from (3-17) to (3-21) in Section 3.1, the first-order derivatives and the Hessian matrix composed of the second-order derivatives with respect to calibration parameters can be obtained as

$$
\nabla L\left(\mu_{x_\mathrm{u}}, s_{x_\mathrm{u}}\right) = \begin{bmatrix} \frac{\partial f}{\partial \mu_{x_\mathrm{u}}} \\ \frac{\partial f}{\partial s_{x_\mathrm{u}}} \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial \mu_y} \times a_\mathrm{u} \\ \frac{\partial f}{\partial s_y} \times {a_\mathrm{u}}^2 \end{bmatrix}
= \begin{bmatrix} -\frac{a_\mathrm{u}}{s_y}\sum_{i=1}^{n}(y_i - \mu_y)) \\ -\frac{{a_\mathrm{u}}^2}{2{s_y}^2}\sum_{i=1}^{n}\left(y_i - \mu_y\right)^2 + \frac{n{a_\mathrm{u}}^2}{2s_y} \end{bmatrix} \tag{3-29}
$$

$$
\mathbf{H}\left(L\left(\mu_{x_\mathrm{u}}, s_{x_\mathrm{u}}\right)\right) = \begin{pmatrix} \frac{\partial^2 f}{\partial \mu_{x_\mathrm{u}}^2} & \frac{\partial^2 f}{\partial \mu_{x_\mathrm{u}} \partial s_{x_\mathrm{u}}} \\ \frac{\partial^2 f}{\partial \mu_{x_\mathrm{u}} \partial s_{x_\mathrm{u}}} & \frac{\partial^2 f}{\partial s_{x_\mathrm{u}}^2} \end{pmatrix}
= \begin{pmatrix} \frac{\partial^2 f}{\partial \mu_y^2} \times {a_\mathrm{u}}^2 & \frac{\partial^2 f}{\partial \mu_y \partial s_y} \times {a_\mathrm{u}}^3 \\ \frac{\partial^2 f}{\partial \mu_y \partial s_y} \times {a_\mathrm{u}}^3 & \frac{\partial^2 f}{\partial s_y^2} \times {a_\mathrm{u}}^4 \end{pmatrix}
= \begin{pmatrix} \frac{n{a_\mathrm{u}}^2}{s_y} & \frac{{a_\mathrm{u}}^3}{{s_y}^2}\sum_{i=1}^{n}(y_i - \mu_y) \\ \frac{{a_\mathrm{u}}^3}{{s_y}^2}\sum_{i=1}^{n}(y_i - \mu_y) & \frac{{a_\mathrm{u}}^4}{{s_y}^3}\sum_{i=1}^{n}\left(y_i - \mu_y\right)^2 - \frac{n{a_\mathrm{u}}^4}{2{s_y}^2} \end{pmatrix} \tag{3-30}
$$

The determinant of the Hessian matrix is then,

$$\det \mathbf{H}\left(L(\mu_{x_\mathrm{u}},\ s_{x_\mathrm{u}})\right) = \frac{n{a_\mathrm{u}}^6}{s_y}\left(\frac{1}{{s_y}^3}\sum_{i=1}^{n}(y_i-\mu_y)^2-\frac{n}{2{s_y}^2}\right)-\frac{{a_\mathrm{u}}^6}{{s_y}^4}\left(\sum_{i=1}^{n}(y_i-\mu_y)\right)^2 \tag{3-31}$$

For a data set of observations ($\mathbf{Y}_{\mathrm{obs}} = \{y_1, \ldots, y_n\}$), let

$$\mu_{y_0} = \frac{1}{n}\sum_{i=1}^{n} y_i,\ \ s_{y_0} = \frac{1}{n}\sum_{i=1}^{n}(y_i-\mu_{y_0})^2 \tag{3-32}$$

Then

$$\frac{1}{n}\sum_{i=1}^{n}(y_i-\mu_y)^2 = s_{y_0} + (\mu_{y_0}-\mu_y)^2 \tag{3-33}$$

$$\sum_{i=1}^{n}(y_i-\mu_y) = \sum_{i=1}^{n} y_i - n\mu_y = n(\mu_{y_0}-\mu_y) \tag{3-34}$$

Thus, the Hessian matrix and the determinant of the Hessian matrix can be reformulated as

$$\mathbf{H}\left(L(\mu_{x_\mathrm{u}},\ s_{x_\mathrm{u}})\right) = \begin{pmatrix} \frac{n}{s_y}{a_\mathrm{u}}^2 & \frac{n(\mu_{y_0}-\mu_y)}{{s_y}^2}{a_\mathrm{u}}^3 \\ \frac{n(\mu_{y_0}-\mu_y)}{{s_y}^2}{a_\mathrm{u}}^3 & \frac{n{a_\mathrm{u}}^4}{{s_y}^3}\left(s_{y_0}+(\mu_{y_0}-\mu_y)^2\right)-\frac{n{a_\mathrm{u}}^4}{2{s_y}^2} \end{pmatrix} \tag{3-35}$$

$$\det \mathbf{H}\left(L(\mu_{x_\mathrm{u}},\ s_{x_\mathrm{u}})\right) = [\frac{n^2}{{s_y}^4}\left(s_{y_0}+(\mu_{y_0}-\mu_y)^2\right)-\frac{n^2}{2{s_y}^3} - \frac{n^2}{{s_y}^4}(\mu_{y_0}-\mu_y)^2]{a_\mathrm{u}}^6 = \frac{n^2}{{s_y}^3}(\frac{s_{y_0}}{s_y}-\frac{1}{2}){a_\mathrm{u}}^6 \tag{3-36}$$

Since $n^2/{s_y}^3$ always results in a positive value, the likelihood function is convex for s_y, satisfying

$$2s_{y_0} > s_y \tag{3-37}$$

3.3.1.2. Analytical Sensitivity Information of the Probability Residual

For convenience of investigation, this section lets $p(y)$ and $q(y)$ be the observed PDF and the predicted PDF of the system response (y), respectively, in Equation (3-25). With the aforementioned assumptions, the probability residual in Equation (3-25) can be reformulated as

$$\begin{aligned}&\int_{-\infty}^{\infty}(f(y)-g(y))^2dy\\&=\frac{1}{\sqrt{4\pi{\sigma_y}^2}}+\frac{1}{\sqrt{4\pi{\sigma_{y_0}}^2}}\\&-\frac{2}{\sqrt{2\pi}}\frac{1}{\sqrt{{\sigma_y}^2+{\sigma_{y_0}}^2}}\mathrm{e}^{-\frac{0.5(\mu_y-\mu_{y_0})^2}{{\sigma_y}^2+{\sigma_{y_0}}^2}}\end{aligned} \tag{3-38}$$

where the statistical parameters ($\boldsymbol{\theta}_{y_0}$) describe the observed PDF characterized by the moments method for a data set of observations ($\mathbf{Y}_{\mathrm{obs}}$) with the previously introduced Gaussian assumption. (For the details of the derivations, please refer to (Lee et al. 2018).) Let z = - $0.5(\mu_y-\mu_{y_0})^2/({\sigma_y}^2+{\sigma_{y_0}}^2)$, a = $1/({\sigma_y}^2+{\sigma_{y_0}}^2)$, b = $\mu_y-\mu_{y_0}$, t = $(2/\pi)^{0.5}e^z a^{3/2}$. The first derivative (gradient vector) of the PR with respect to two statistical parameters, mean (μ) and standard deviation (σ) of the predicted PDF, is represented as follows

$$\begin{bmatrix}\frac{\partial PR}{\partial\mu_y}\\\frac{\partial PR}{\partial\sigma_y}\end{bmatrix}=\begin{bmatrix}\frac{\sqrt{\frac{2}{\pi}}e^z(\mu_y-\mu_{y_0})}{({\sigma_y}^2+{\sigma_{y_0}}^2)^{3/2}}\\-\frac{1}{2\sqrt{\pi}{\sigma_y}^2}-\frac{\sqrt{\frac{2}{\pi}}e^z(\mu_y-\mu_{y_0})^2\sigma_y}{({\sigma_y}^2+{\sigma_{y_0}}^2)^{5/2}}+\frac{e^z\sqrt{\frac{2}{\pi}}\sigma_y}{({\sigma_y}^2+{\sigma_{y_0}}^2)^{3/2}}\end{bmatrix} \tag{3-39}$$

Differentiating the gradient vector once again, a matrix of the second partial derivatives for the probability residual (PR) can be obtained as

$$\begin{pmatrix} \frac{\partial^2 PR}{\partial \mu_y^2} & \frac{\partial^2 PR}{\partial \mu_y \partial \sigma_y} \\ \frac{\partial^2 PR}{\partial \sigma_y \partial \mu_y} & \frac{\partial^2 PR}{\partial \sigma_y^2} \end{pmatrix}$$

where

$$\frac{\partial^2 PR}{\partial \mu_y^2} = t(1 - ab^2) \tag{3-40}$$

$$\frac{\partial^2 PR}{\partial \mu_y \partial \sigma_y} = tab\sigma_y(-3 + ab^2)$$

$$\frac{\partial^2 PR}{\partial \sigma_y^2} = \frac{1}{\sqrt{\pi}{\sigma_y}^3} - t(a^2 b^2 {\sigma_y}^2 (ab^2 - 3) - (3a{\sigma_y}^2 - 1)(ab^2 - 1)$$

Using the Equations from (3-17) to (3-21) in Section 3.1, the first-order derivatives and the Hessian matrix composed of the second-order derivatives with respect to calibration parameters can be obtained as

$$\mathbf{H}\left(PR\left(\mu_{x_\mathrm{u}},\ \sigma_{x_\mathrm{u}}\right)\right) = \begin{pmatrix} \frac{\partial^2 PR}{\partial \mu_y^2} \times {a_\mathrm{u}}^2 & \frac{\partial^2 PR}{\partial \mu_y \partial \sigma_y} \times |a_\mathrm{u}| \times a_\mathrm{u} \\ \frac{\partial^2 PR}{\partial \mu_y \partial \sigma_y} \times |a_\mathrm{u}| \times a_\mathrm{u} & \frac{\partial^2 PR}{\partial \sigma_y^2} \times {a_\mathrm{u}}^2 \end{pmatrix} \tag{3-41}$$

The determinant of the Hessian matrix is then

$$\det \mathbf{H}\left(PR\left(\mu_{x_\mathrm{u}}, \sigma_{x_\mathrm{u}}\right)\right) = \left[t(1 - ab^2)\frac{1}{\sqrt{\pi}{\sigma_y}^3} - t^2\left(a{\sigma_y}^2(a^2 b^4 + 3) - (ab^2 - 1)^2\right)\right] \times {a_\mathrm{u}}^4 \tag{3-42}$$

Finally, the probability residual is convex for μ_{x_u} and σ_{x_u}, satisfying

$$t(1 - ab^2)\frac{1}{\sqrt{\pi}{\sigma_y}^3} - t^2\left(a{\sigma_y}^2(a^2 b^4 + 3) - (ab^2 - 1)^2\right) > 0 \tag{3-43}$$

Through examination, it can be seen that OBMC with two calibration metrics is not convex for the entire design domain, because the determinants of Hessian matrices in Equations (3-36) and (3-42) are not always positive for the entire design domain. The entire design domain is separated into convex and concave domains.

The boundaries between the convex and concave domains are called inflection lines. Thus, in conclusion, OBMC using the existing two calibration metrics cannot take advantage of the benefits of convexity.

3.3.1.3. The Global Optimum of OBMC Using Calibration Metrics

In the previous section, it is confirmed that the optimization problem of OBMC with two calibration metrics is not globally convex. To further understand the causes of the unstable and inaccurate calibrated results introduced in Figure 1-5, this section first obtains the point that satisfies the first-order necessary condition for the optimum of OBMC. The point satisfying the first-order necessary condition for the optimum (($\mu_{L,\text{ opt}}$, $\sigma_{L,\text{ opt}}$) denotes the optimum of OBMC with the likelihood function (L); ($\mu_{PR,\text{ opt}}$, $\sigma_{PR,\text{ opt}}$) denotes the optimum of OBMC with the probability residual (PR)) and can be obtained by letting the first derivatives of the calibration metrics equal zero, as shown:

$$\nabla L(\mu_{L,\text{opt}}, s_{L,\text{opt}}) = \begin{bmatrix} -\frac{1}{s_{L,\text{opt}}}\sum_{i=1}^{n}(y_i - \mu_{L,\text{opt}}) \times a_{\text{u}} \\ (-\frac{1}{2s_{L,\text{opt}}{}^2}\sum_{i=1}^{n}\left(y_i - \mu_{L,\text{opt}}\right)^2 + \frac{n}{2s_{L,\text{opt}}}) \times a_{\text{u}}{}^2 \end{bmatrix} = 0 \tag{3-44}$$

$$\begin{bmatrix} \mu_{L,\text{opt}} \\ s_{L,\text{opt}} \end{bmatrix} = \begin{bmatrix} \frac{\sum_{i=1}^{n} y_i}{n} \\ \frac{\sum_{i=1}^{n}\left(y_i - \mu_{L,\text{opt}}\right)^2}{n} \end{bmatrix} = \begin{bmatrix} \mu_{y_0} \\ s_{y_0} \end{bmatrix} \tag{3-45}$$

$$\nabla L(\mu_{PR,\text{opt}}, s_{PR,\text{opt}}) = 0 \tag{3-46}$$

$$\begin{bmatrix} \mu_{PR,\text{opt}} \\ \sigma_{PR,\text{opt}} \end{bmatrix} = \begin{bmatrix} \mu_{y_0} \\ \sigma_{y_0} \end{bmatrix} \tag{3-47}$$

The point presented in Equations (3-45) and (3-47) is not necessarily the optimum; thus, it should be checked whether the point satisfies the second-order sufficient condition. By substituting the point ($\mu_{L,\text{ opt}}$, $\sigma_{L,\text{ opt}}$) and ($\mu_{PR,\text{ opt}}$, $\sigma_{PR,\text{ opt}}$) into Equations (3-36) and (3-42), respectively, it can be checked whether the determinant of the Hessian matrix is positive definite. In conclusion, the points ($\mu_{L,\text{ opt}}$, $\sigma_{L,\text{ opt}}$) and ($\mu_{PR,\text{ opt}}$, $\sigma_{PR,\text{ opt}}$) in Equations (3-45) and (3-47) are indeed the optimum of OBMC performed with the two calibration metrics.

The conclusion of this section shows that a single optimum, which is the global optimum, exists in the optimization problem that is solved using either of the two calibration metrics. In other words, using either of the two calibration metrics, the issue of OBMC converging to a local optimum is not a concern anymore. (It is confirmed that there is no local optimum.) Therefore, the phenomenon of OBMC stopping at inaccurate points can be explained by the existence of a large flat space associated with semi-definite Hessian matrices, where the optimization algorithms cannot find the proper direction that would lead to the global optimum.

3.3.2 Numerical Investigation

To follow up on the previous section, numerical investigation of OBMC with two calibration metrics is conducted to help understand the analytical results from previous section using a simple numerical example. For example, the linear model in Equation (3-12) has three uncertain input variables (**X**, two known and one unknown) and one system response (y). The values for model parameters a_1, a_2, a_3, and b are chosen as 0.5, -0.5, 1, and 0, respectively. The two known input variables follow normal distributions ($x_{1,\text{known}}$ ~ $N(0,1)$, $x_{2,\text{known}}$ ~ $N(0,1)$). A model calibration problem is formulated to estimate the statistical parameters of an unknown input variable ($x_{3,\text{unknown}}$) that is assumed to follow a normal distribution. For an arbitrary normalized data set of observations ($\mathbf{Y}_{\text{obs}} = (y_1, y_2, ..., y_n)$), such as $\mathbf{Y}_{\text{obs}} = (-1, 0, 1)$, the true mean ($\mu_0$) and standard deviation (σ_0) of the unknown input variable are assigned as '0' and '1' ($s_0 = 1$), respectively, when the method of moments is used to estimate the statistical moments of the population.

Figure 3-2 shows the 2- and 3-dimensional response surfaces of the two calibration metrics in terms of variations in the two calibration parameters (mean and standard deviation) of the unknown input ($x_{3,\text{ unknown}}$). The exact same optimum point, (μ_0, σ_0) = (0, 1), is shown as a point on the two response surfaces provided by the two calibration metrics. As confirmed in the analytical examination, optimization problems using the two metrics are not convex over the entire design domain. Figure 3-2b and Figure 3-2d, which are seen from an aerial view (i.e., 2-dimensional plot) of Figure 3-2a and Figure 3-2c, respectively, show that each response surface is separated into convex (the red shaded area) and concave (the blue shaded area) domains by an inflection line (a yellow dotted line). The global optimum of this example is identified as the points within the convex domains. It is shown that the inflection line of the likelihood function is a straight line, and a relatively large convex domain, $\mu \in R$ and $\sigma \leq 2$, is clearly separated from the concave domain.

It is important to notice that a part of the response surface is flat. For further examination, Figure 3-3a and Figure 3-3b show, respectively for each metric, the determinant of the Hessian matrix of the objective function with respect to the two calibration parameters (mean and standard deviation) of the unknown input. A large portion of the determinant surface of the Hessian matrix between clear convex and concave areas have values close to zero; this means there is little change in slopes of the objective function with respect to the two calibration parameters of the unknown input. This means that the optimization may not progress toward the optimum point unless the exact gradient information is provided. For example, if there is no given gradient information, a gradient optimization algorithm can approximately calculate the gradient at the current point using finite difference gradient calculations; this may cause the issues depicted in Figure 1-5. This concern can be solved when the optimization is conducted using explicit sensitivity information, which is available from derivative derivations for the convexity checking. In addition, an extremely low degree of slope and approximate calculation of sensitivity information can be exacerbated by the randomness that arises due to the sampling method, if it is used for the uncertainty propagation process. In this case, a large enough number of samples is required to produce a robust and accurate result of uncertainty propagations. Otherwise, other uncertainty propagation methods can be devised to solve this problem. Discussion of uncertainty propagation methods is continued in next chapter.

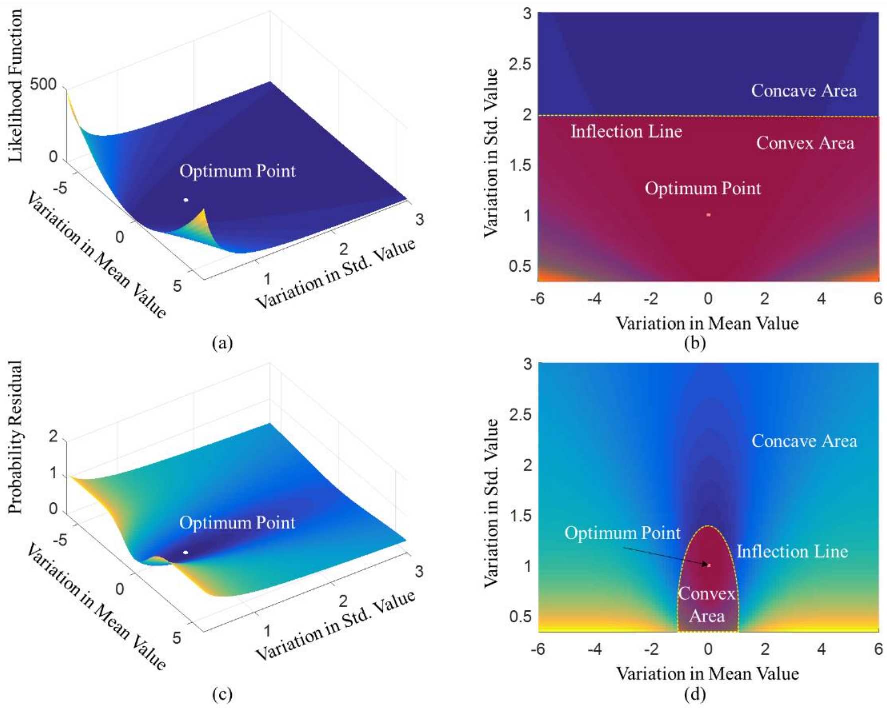

Figure 3-2 Response Surface (RS) of Calibration Metrics for the Calibration Parameters (the Mean and Standard Deviation of the Unknown Input Variable): (a) 3-D RS of the Likelihood Function, (b) 2-D RS of the Likelihood Function, (c) 3-D RS of the PR, and (d) 2-D RS of the PR

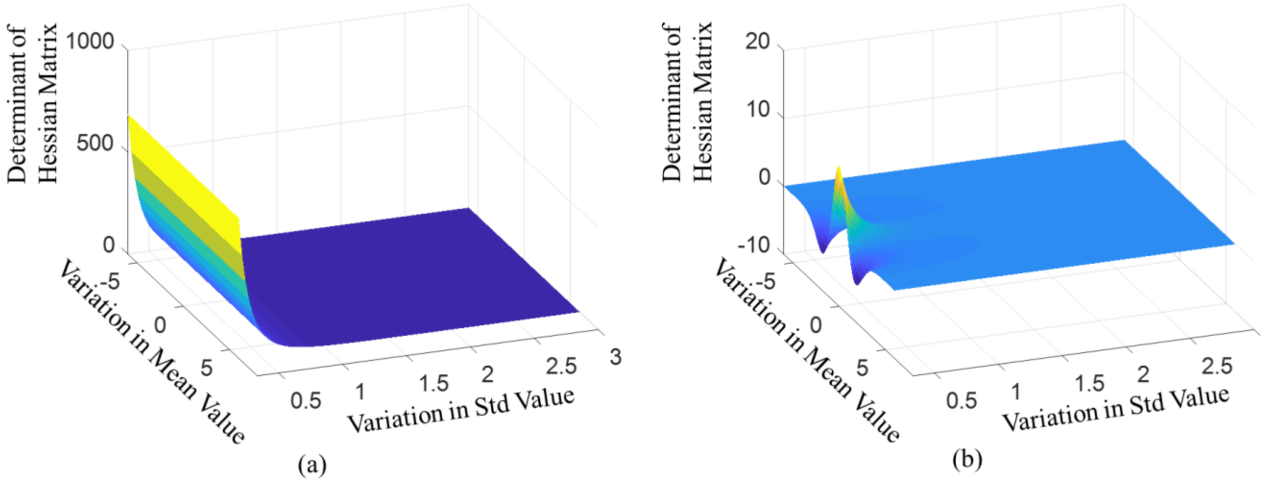

Figure 3-3 Determinant of the Hessian Matrix for the Calibration Parameters (the Mean and Standard Deviation of the Unknown Input Variable), as Derived Using: (a) the Likelihood Function, (b) the Probability Residual

3.4 Robust Optimization-Based Model Calibration with Analytical Sensitivity Information

As previously mentioned, gradient-based algorithms are recommended for the optimization algorithms used in OBMC due to their efficiency. Gradient-based algorithms for unconstrained gradient-based optimization can be described as shown in Table 3-1. In summary, two major issues related to the use of gradient-based algorithms are 1) computing the search direction ($\mathbf{p}_k$) and 2) finding the step size (α_k). To find the step size, the search direction should be determined first.

Success of optimization with gradient-based algorithms (e.g., steepest descent method, Conjugate gradient method, Newton's method, modified Newton method, quasi-Newton method, trust region methods, or sequential quadratic programming), requires an exact calculation of sensitivity information, including the first-order, second-order, and Hessian. (Note that if an optimization is globally convex, it is guaranteed to converge to the global optimum with various advanced gradient-based algorithms. However, as shown in the previous section, it is proved that OBMC performed with the two existing calibration metrics is not globally convex.)

For exact calculation of sensitivity information, analytical sensitivity information is desired. However, analytical sensitivity information is not readily available for most optimization problems. In this case, finite difference gradient calculations (Venter 2010) provide approximated gradient information. Unfortunately, the accuracy of finite difference calculations depends on the selected step size, which means the optimization may not be robust unless a proper step size is used. Furthermore, estimating the gradient information using finite difference calculations requires expensive computational time due to the required function evaluations.

Fortunately, if the analytical gradient information is provided, then gradient-based optimization can produce more efficient and accurate optimized results. Thus, this study attempts to use the analytical gradient information, which is derived during the examination of global convexity in Section 3.3.1, to improve the results of the optimization-based model calibration problems. The effectiveness of calibration by OBMC with analytical sensitivity information is checked in the case study section.

Table 3-1 General Steps for Unconstrained Gradient-Based Optimization Algorithms

Step 1: Start optimization	Start optimization with iteration number $k = 0$ and a starting point, $\mathbf{x}_k$.
Step 2: Test for convergence	If the conditions for convergence are satisfied, then the optimization can stop and $\mathbf{x}_k$ is the solution.
Step 3: Compute a search direction	Compute the vector $\mathbf{p}_k$ that defines the direction in a space along which optimization progresses.
Step 4: Compute the step length	Find a positive scalar, α_k, such that $f(\mathbf{x}_k + \alpha_k\mathbf{p}_k) < f(\mathbf{x}_k)$.
Step 5: Update the design variables	Set $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k\mathbf{p}_k$, $k = k + 1$ and go back to Step 1.

3.5 Summary and Discussion

The motivation of Research Thrust 2 was to address the unstable and inaccurate calibrated results, which implies that the optimization of OBMC with existing calibration metrics was not terminated at the global optimum. Analytical and numerical investigations of optimization for OBMC with the two existing calibration metrics have confirmed that 1) OBMC with the existing calibration metrics is not globally convex. In other words, many gradient-based algorithms, which have advantages for searching for the global optimum, cannot work for OBMC when used with the two existing calibration metrics; 2) Under some assumptions, no local minima exists, however, this is not guaranteed for other cases, such as the nonlinear model or other types of probability distribution; 3) A large portion of the determinant surface of the Hessian matrix has a near zero value, which may cause difficulty in obtaining the sensitivity information, which includes the first- and second-order derivatives, and Hessian. In the end, this may cause the optimization to not progress toward the expected optimum point; 4) Even a small degree of randomness due to sampling methods for the uncertainty propagation process can mean that the optimization does not progress toward the global optimum, when it is related to the third issue. Four conclusions of analytical and numerical investigations are summarized in Figure 3-4. To overcome this situation, the analytical results were used for the sensitivity information of OBMC. To prove the effectiveness of the proposed idea, a case study of a "bearing capacity equation" (introduced in Section 6.2) is examined to show that OBMC with analytical sensitivity information provides better calibrated results, compared with the sensitivity information calculated through a finite difference approach embedded in the optimization solvers.

Confirm 1: OBMC is Not Globally Convex.	Confirm 2: No Local Optima
• Analytical Investigation: Likelihood Function: Not Globally Convex Probability Residual: Not Globally Convex	• Analytical Investigation: Likelihood Function: No Local Optimum Probability Residual: No Local Optimum
→ Unable to use convex optimization methods	→ No need for global optimization

Confirm 3/4: Low Degree of Sensitivity & Approximate Calculation of Sensitivity & Randomness of Sampling Methods

- Low Degree of Sensitivity → Searching the optimization direction is very sensitive to the accuracy of calculation of sensitivity information.
- Approximate Sensitivity → Calculations of sensitivity by finite difference gradient leads to searching wrong optimization direction.
- Randomness of Sampling Methods → Randomness by not enough numbers of samplings boosts wrong optimization direction

→ Accurate and robust OBMC is available by providing analytical information.

Figure 3-4 The Results of Analytical and Numerical Investigation on Possible Causes of Inaccurate and Unstable Calibrated Results

One limitation of this study is that four assumptions are required for deriving the analytical sensitivity information. Assumptions include: 1) the model is in a linear relationship, 2) only one unknown input variable exists in the model, 3) the uncorrelated input variables of the model follow a normal distribution, and 4) the simulation and experimental data both follow normal distributions. For example, if a model is in a nonlinear relationship and the variability of a system response does not follow a normal distribution, then the obtained analytical information is not applicable.

To expand use of the proposed method, it will be possible to devise a new calibration metric that constructs a convex optimization problem for OBMC. This inspires an idea for next research thrust, described in Chapter 4. In addition, limitations due to the assumptions about the type of probability distribution of a system response can be overcome by a follow-up study, such as a study on deriving gradient information for different types of PDFs of a system response.

Sections of this chapter have been published or submitted as the following journal article:

1) Guesuk Lee, Guilian Yi, and Byeng D. Youn, "A Comprehensive Study on Enhanced Optimization-Based Model Calibration Using Gradient Information," *Structural and Multidisciplinary Optimization*, Vol. 57, No. 5, 2018.

Chapter 4

Sequential Optimization and Uncertainty Propagation Method for Efficient Optimization-Based Model Calibration

Optimization-based model calibration (OBMC) adopts optimization under uncertainty (OUU) to calibrate statistical parameters (e.g., the mean and standard deviation) of the unknown input variables of the computational model. OUU associates the optimization process with the uncertainty propagation (UP) process (also called probabilistic assessment) where uncertainties in input variables are propagated through a model to obtain the variability in the system responses (Eldred et al. 2002) (Swiler et al. 2008) (Youn et al. 2011) (Yao et al. 2011) (Arendt et al. 2012a) (Fender et al. 2014) (De Cursi and Sampaio 2015) (Jung et al. 2016) (Lee et al. 2018) (Hu et al. 2018). The computational cost of OUU associating the optimization and UP processes in a loop is highly expensive. In summary, formulating an OUU for OBMC and selecting a method for the UP process are important issues for developing efficient OBMC. The objective of Research Thrust 3 is thus to develop an efficient OUU formulation for OBMC, while retaining the accuracy. The new OUU formulation is based on a comprehensive study of the OUU formulations and UP methods developed from the society of reliability-based design optimization (RBDO).

First, Research Thrust 3 begins with a review of studies on the developed OUU formulations and UP methods. The field of RBDO, which is a field of study adopting OUU formulations, has developed various OUU formulations (e.g., double-loop, single-loop, decoupled RBDO formulations, and others (Enevoldsen and Sørensen 1994) (Tu et al. 1999) (Du and Chen 2002) (Youn et al. 2004) (Aoues and Chateauneuf 2010)) and reliability assessment methods (e.g., expansion method, most probable point method, sampling method, approximate integration method, and others (Li et al. 2003) (Rahman and Xu 2004) (Xu and Rahman 2004) (Lee et al. 2008a) (Lee et al. 2008b) (Youn et al. 2008)). (Note that the reliability assessment in RBDO refers to the UP process in OBMC, also called probabilistic assessment or probabilistic analysis.) Research Thrust 3 thus elects an appropriate OUU formulation and a UP method for developing an efficient and accurate OBMC.

Next, Research Thrust 3 inaugurates a new formulation for OBMC called sequential optimization and uncertainty propagation (SOUP). The proposed SOUP employs a two-stage, sequential, single-loop OUU for an efficient and accurate OBMC. The first stage of SOUP is formulated to utilize an *efficient* process. Then, the second stage of SOUP formulates an *accurate* process. Thereby, SOUP aims to achieve a process for OBMC that is both efficient and accurate. Meanwhile, two calibration metrics, the moment matching metric and probability residual, are devised to be used in each sequence of SOUP.

The remainder of Chapter 4 is organized as follows. Sections 4.1 and 4.2, respectively, review OUU formulations and UP methods developed from the field of RBDO. In Section 4.3, the proposed formulation for an efficient and accurate OBMC is introduced, with discussion of two UP methods and two calibration metrics. The effectiveness of SOUP is investigated through examination of a case study "bearing capacity equation" presented later in Section 6.2. Finally, the conclusions of Research Thrust 3 are provided in Section 4.4.

4.1 Brief Review of Optimization Under Uncertainty Formulations

First, Section 4.1.1 provides an overview of four types of OUU formulations. Section 4.1.2 provides a summary and discussion of the possible use of existing OUU formulations for OBMC.

4.1.1 Overview of Optimization Under Uncertainty Formulations

The society of reliability-based design optimization (RBDO) has developed various OUU formulations to achieve a reliable design by associating the design optimization and reliability assessment processes. Based on the author's intention, developed OUU formulations can be categorized into four types: 1) basic single-loop, 2) double-loop, 3) single-loop, and 4) decoupled formulations (Du and Chen 2002) (Aoues and Chateauneuf 2010) (Yoon 2018). Table 4-1 summarizes the four categories OUU formulations and algorithm examples.

Basic single-loop RBDO: A basic way to formulate an OUU for RBDO is to organize a single-loop RBDO where the single loop includes the reliability assessment process, as depicted in Figure 4-1a. Normally, Monte Carlo simulations (MCS) are used for the sampling-based reliability assessment process for a single-loop RBDO. The sampling methods ensure high accuracy when a sufficient number of samples are available. However, sequential execution of the two processes is too computationally expensive to be applied to real-world applications. Note that other reliability assessment methods can substitute MCS for efficiency, but with lower accuracy. A review of the reliability assessment methods, which refer UP methods for OBMC, is presented in Section 4.2.

Double-loop RBDO: Double-loop formulations (e.g., the reliability index approach-based RBDO (Nikolaidis and Burdisso 1988) and the performance measure approach-based RBDO (Tu et al. 1999)) were developed to substitute for expensive sampling methods in a single-loop RBDO (basic single loop RBDO) by formulating nested optimization problems, where the inner loop deals with the

Table 4-1 Summary of Optimization Under Uncertainty Formulations

OUU Formulation	**Algorithm Example**
Basic single-loop RBDO	Sampling method-based RBDO
Double-loop RBDO	Reliability index approach-based RBDO Performance measure approach-based RBDO
Single-loop RBDO	Single-loop single vector
Decoupled RBDO	Sequential optimization and reliability assessment

reliability assessment (UP) and the outer loop deals with the design optimization (Figure 4-1b). The fundamental idea of the double-loop formulation is to conduct the reliability assessment process by measuring the direct distance (reliability index) from the design point to the constraints. However, the computational cost still can be high because the inner loop formulates another optimization inside of the outer loop to find the optimal point on the constraints to calculate the reliability.

Single-loop and Decoupled RBDO: To ease the computational cost by separating the double-loop, single-loop approaches (e.g., single-loop, single-vector (Liang et al. 2008) (Nguyen et al. 2010)) and decoupled approaches (e.g., sequential optimization and reliability assessment (Du and Chen 2002)) have been developed. Single-loop RBDO eliminates the inner loop for the reliability assessment by approximating probabilistic constraints to deterministic ones (Figure 4-1c). Probabilistic constraints are approximated into deterministic ones and then simple design optimization is conducted without additional reliability assessment. Decoupled RBDO decouples two loops into the outer loop for deterministic design optimization and the inner loop for the reliability assessment (Figure 4-1d). The two separated loops are performed sequentially until a design optimization converges. Compared to the double-loop RBDO, which conducts the reliability assessment for all design changes in the outer loop, the decoupled RBDO conducts the reliability assessment only once after the deterministic optimum design from the outer loop is achieved.

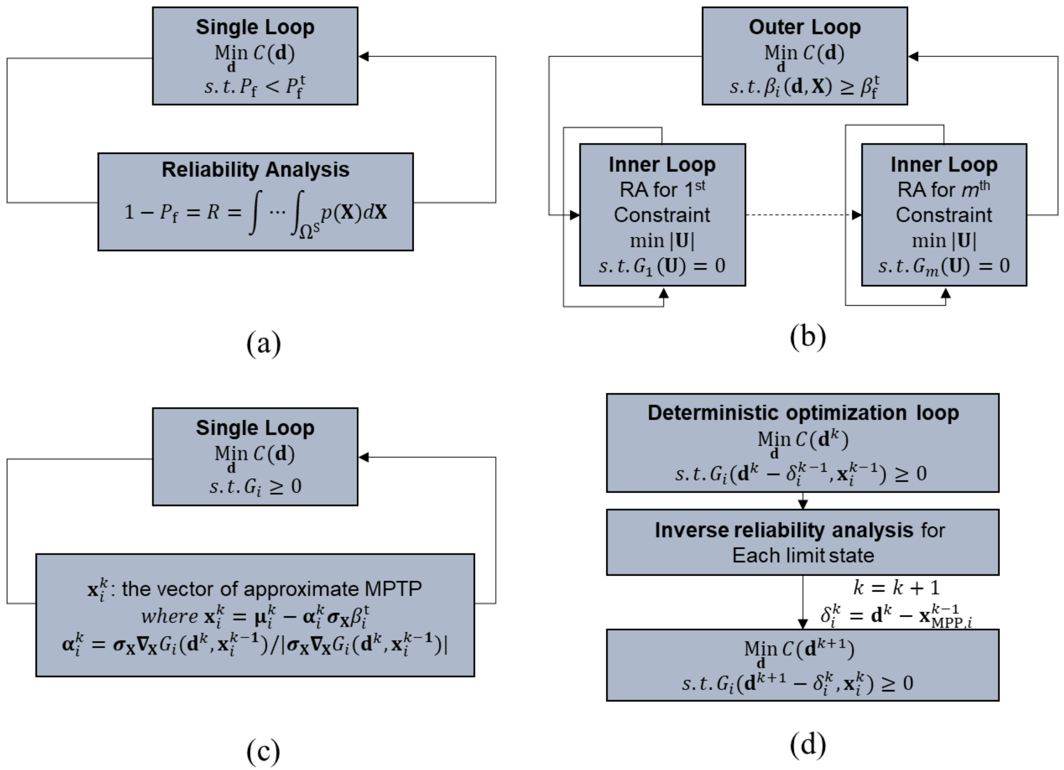

Figure 4-1 Formulations of Optimization Under Uncertainty: a) Basic Single-Loop, b) Double-Loop, c) Single-Loop, and d) Decoupled Reliability-Based Design Optimization

4.1.2 Summary and Discussion

To summarize the review of the developed RBDO formulations, RBDO formulations, such as double loop, single loop, or decoupled RBDO, were available for two reasons: 1) the reliability assessment process is performed with respect to the constraints, 2) the result of the reliability assessment process can be conservatively approximated. However, OBMC is an unconstrained optimization problem. Therefore, OBMC cannot take the advantages of the OUU formulation associating reliability constraints. Furthermore, the result of the UP should be accurate, because a probability distribution describing the variability in a system response should be accurate enough to be compared with the one from experimental observations. In conclusion, OBMC can only use a basic single-loop formulation. Thus, the way to increase the efficiency of OBMC is to either take advantage of the basic single loop formulation, or to increase the efficiency of the UP process within it.

4.2 Brief Review of Uncertainty Propagation Methods

The conclusion of the review in Section 4.1 implies that there is still a limitation that can be addressed to improve the efficiency of OBMC by formulating an OUU developed from the society of RBDO. The remaining way to improve the efficiency of OBMC is to adopt a proper UP method, which can be used in the basic single-loop formulation. This section begins with a review of uncertainty propagation methods developed to date by members of the RBDO society (Section 4.2.1). After that, a summary and discussion about electing a proper UP method for OBMC is provided in Section 4.2.2.

4.2.1 Overview of Uncertainty Propagation Methods

Various techniques for UP methods have been developed through various engineering societies. Existing UP methods can be broadly classified into three categories: 1) analytical methods, 2) sampling methods, and 3) numerical integration methods (Fonseca et al. 2002).

4.2.1.1. Analytical Methods

Based on the author's understanding, analytical methods for UP, such as expansion methods and most probable point methods, are considered inadequate or inapplicable for the UP process of OBMC. Expansion methods are simple to implement, but require partial derivatives of the system responses and are inaccurate for non-Gaussian cases (Schuëller and Jensen 2008). The method of most probable point is efficient and accurate for reliability-based design optimization (Lee et al. 2008c) (Lee et al. 2010). However, the reliability index, which is the result of the most probable point method, is inappropriate for OBMC, where the probabilistic characteristics of the system responses are required.

4.2.1.2. Sampling Methods

Basically, UP is designed to determine the probabilistic characteristics (e.g., a probability distribution or statistical moments) of the system responses. For OBMC, where probability distributions from computational predictions and experimental observations are quantitatively compared through a calibration metric, a probability distribution is the most desirable way to describe the probabilistic characteristics of the system responses. In this respect, sampling methods, of which result in a full description of uncertainty in the system response by a probability distribution, could be an appropriate method.

Sampling methods (e.g., direct-, quasi- Monte Carlo simulation, importance sampling) generate independent samples of all input variables of a computational model and repeat deterministic simulations to obtain the probability distributions of the system responses (Rahman and Xu 2004) (Beyer and Sendhoff 2007) (Yao et al. 2011). To obtain a fully accurate probability distribution of the system responses, however, a large number of simulations is required. Thus, sampling methods are considered impractical for the OUU where the UP is incorporated with the optimization process. The computational burden can be serious when the computational prediction includes expensive calculations (e.g., a system-level finite-element model). In general, sampling methods are used for benchmark studies where accurate computation is required.

4.2.1.3. Numerical Integration Methods

Statistical moments are another way to describe the probabilistic characteristics of the system responses. Calculating the statistical moments of the system responses entails numerical integration. For an N-variable computational model, however, N-dimensional integration, of which computational cost exponentially increases as the number of input variables increases, is required to calculate the statistical moments. Approximate integration methods, such as the univariate dimension reduction (UDR) method (Rahman and Xu 2004) (Lee et al. 2008b), the generalized dimension reduction method (Xu and Rahman 2004), and the eigenvector dimension reduction method (Youn et al. 2008) (Youn and Wang 2008), have been proposed and developed to efficiently calculate the statistical moments by simplifying one multi-dimensional integration into multiple one-dimensional integrations by using additive decomposition.

Approximation integration methods have strengths in terms of computational cost. Their computational costs only increase additively (rather than exponentially) when the number of random variables increases. However, two issues exist that limit the use of approximate integration methods in OBMC. First, the final result of an approximate integration method is a set of statistical moments. With a few statistical moments, approximate comparison is available; for example, for the first four statistical moments (mean, standard deviation, skewness, and kurtosis), probability distributions of the system responses can be approximately constructed by the Pearson system (Youn and Wang 2008) (Youn and Xi 2009) (Choi et al. 2010). However, accurate comparison between two probability distributions from experimental observations and computational predictions is not available. The second issue is the numerical error that arises due to high non-linearity or large random variations. The numerical error from using univariate dimension reduction or eigenvector dimension reduction is less than the second-order Taylor series approximation. (It is noteworthy that the numerical error is less than that found in analytical methods.) This means that numerical error occurs, for example, when a computational model includes a cross-term, the term including multiplication of two variables. This numerical error is due to additive decompositions eliminating the cross-term in the integration. Although the degree of error is less than the second-order Taylor series approximation, the numerical error may not be trivial for a particular case.

4.2.2 Summary and Discussion

In summary, among three broadly classified UP methods, sampling methods and numerical integration methods are available to be used in OBMC, since their results provide probabilistic characteristics (e.g., a probability distribution derived from a sampling method or statistical moments from numerical integration methods) of the system responses. Comparing these two methods, sampling methods have strength in terms of their high accuracy, and numerical integration methods have strength in efficiency, while retaining a degree of accuracy. The strengths from each method inspire formulation of a new OUU formulation for OBMC.

In this study, the direct Monte Carlo sampling (MCS) method is used to represent the sampling methods and the univariate dimension reduction (UDR) method is used to represent the approximation integration methods. Other advanced methods for each category may show better accuracy and efficiency; however, this study focuses on a representative method for each category that can be used in OBMC without difficulty.

4.3 Sequential Optimization and Uncertainty Propagation for Optimization-Based Model Calibration

This section introduces the proposed OUU formulation for an efficient and accurate OBMC. First, Section 4.3.1 explains how the proposed method – sequential optimization and uncertainty propagation (SOUP) – is formulated for OBMC. Second, Section 4.3.2 explains how the selected UP methods – the direct Monte Carlo simulations (direct MCS) and univariate dimension reduction (UDR) – are adopted in SOUP for OBMC. Third, Section 4.3.3 discusses on the two calibration metrics – the proposed moment-matching metric (MMM) and probability residual (PR).

4.3.1 Scheme of Sequential Optimization and Uncertainty Propagation

From the comprehensive study of OUU formulations and UP methods found in Sections 4.1 and 4.2, it can be concluded that 1) the available OUU formulation for OBMC is the basic single optimization loop, and 2) the available UP methods for OBMC are sampling methods (e.g., direct MCS) and approximation integration methods (e.g., UDR). As discussed in Section 4.2, sampling methods are accurate but computationally expensive. On the other hand, approximation integration methods are efficient, but have questionable accuracy (numerical errors may exist).

To take advantage of the positive aspects of each UP method, the proposed method links two sequential optimization loops that involve both of the UP processes, as shown in Figure 4-2. The proposed method is named sequential optimization and uncertainty propagation (SOUP). The first sequence of SOUP, which is called *the* *efficient sequence*, employs approximate integration methods for the UP process. In this paper, UDR is used to represent the approximate integration methods. As depicted in Figure 4-2, the first sequence of OBMC calibrates the initial calibration parameter ($\boldsymbol{\theta}_{\mathbf{X}\text{u}, k=1}$) with an approximate integration method and moment matching metric. The optimal (final) calibration parameters of the first sequence of OBMC becomes the intermediate calibration parameters ($\hat{\boldsymbol{\theta}}_{\mathbf{X}\text{u}}$), which then are the initial calibration parameters for the second sequence of OBMC. In the second sequence of SOUP, which is called *the accuracy sequence*, the optimization associated with accurate sampling methods aims to search for the optimal values for the calibration parameters ($\hat{\boldsymbol{\theta}}_{\mathbf{X}\text{u, opt}}$). Though the computational efficiency is low when using sampling methods in OBMC, SOUP's efficiency is improved as compared to using OBMC entirely with sampling methods. This is because the initial point of OBMC with sampling methods is closer to the potential optimum; thereby, the function counts until the convergence by the OBMC with sampling methods decreases. Despite the computational cost of the second sequence of SOUP (due to its use of sampling methods), the second sequence of SOUP can guarantee accuracy, since the UP result derived from the sampling methods is a full description of the uncertainty in a system response, through which an accurate calibration is available.

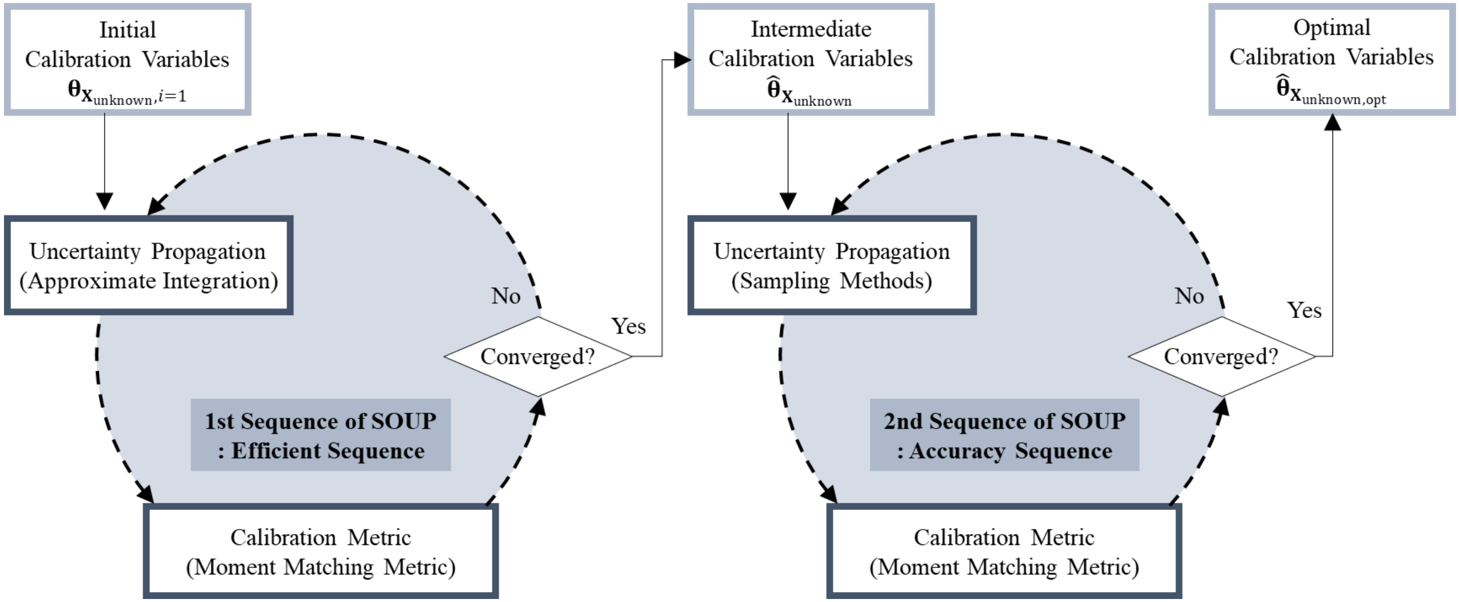

Figure 4-2 Sequential Optimization and Uncertainty Propagation for an Efficient and Accurate Optimization-Based Model Calibration

4.3.2 Uncertainty Propagation Methods in SOUP

For the proposed SOUP procedure, two UP methods, one using a sampling method and the other using an approximate integration method, are used to formulate a two-sequence OUU. In this paper, the direct MCS and univariate dimension reduction method are used to represent the sampling methods and approximate integration methods, respectively. Section 4.3.2.1 introduces a procedure to adopt direct MCS in OBMC and provides a discussion about the accuracy and efficiency issue. Section 4.3.2.2 presents how UDR eases the computational cost in the UP process and the reason why the numerical errors occur.

4.3.2.1. Direct Monte Carlo Simulations

Direct Monte Carlo sampling (MCS) or simulation is one representative method that belongs to the general category referred to as sampling methods. Table 4-2 summarizes the five steps used to conduct direct MCS. Direct MCS repeatedly evaluates a deterministic prediction through a computational model using sets of random numbers as input variables (Step 4). Due to randomness in generating sets of random input variables (Step 2), a sufficient number (m) of repetitions (Step 4) is required to obtain a robust result of the predictions. Herein, robustness in predictions means the convergence of constructed probability distributions or calculated statistics in Step 5.

The strength of sampling methods, including direct MCS, is that the UP result will be accurate if a sufficient number (m) of repeating deterministic predictions from sets of random numbers is promised. To analyze the repeated predictions, probability distributions can be constructed either by parametric methods (e.g.,

Table 4-2 Steps for Direct Monte Carlo Sampling for Optimization-Based Model Calibration

Step 1	Define a computational model ($y = M_{\text{model}}(x_1, x_2, \ldots, x_N)$) with n number of input variables.
Step 2	Generate a set of random input variables ($\mathbf{X}_i = (x_{i1}, x_{i2}, \ldots, x_{iN})$).
Step 3	Evaluate the model (M_{model}) and store the prediction as y_i.
Step 4	Repeat steps 2 and 3 for $i = 1$ to m.
Step 5	Analyze the predictions ($\mathbf{Y}_{\text{pre}} = (y_1, y_2, \ldots, y_m)$) by constructing probability distributions ($p(y)$) or calculating statistics ($\boldsymbol{\theta}_{\mathbf{Y}}$).

maximum likelihood estimation with an assumption of the type of probability distribution) or non-parametric methods (e.g., Kernel density estimation). Because a full description of the probability distribution of the system responses is available with a sufficient number of samples, use of direct MCS in OBMC improves the accuracy of the calibrated results.

As the number (m) of repetitions increases, however, the computational cost for the UP increases. The computational burden gets even worse when the deterministic prediction using one set of random input variables is also expensive (e.g., a complicated finite difference element model of an engineered system). Therefore, it is too expensive to proceed with the entire process of OBMC until the initial part converges to the optimum.

4.3.2.2. The Univariate Dimension Reduction Method

Uncertainty propagation (UP) using numerical integration methods is designed to obtain the statistical moments (e.g., mean, standard deviation, skewness, kurtosis) that describe the uncertainty in the system response (y). To calculate the m^{th} statistical moment of the predicted system response (y_{pre}), a multivariate (N-dimensional) integration is required, as

$$E[y^m(\mathbf{X})] \equiv \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \{y(x_1, \dots, x_N\}^m p_{x_1,\dots,x_N}(x_1, \dots, x_N) dx_1 \cdots dx_N \tag{4-1}$$

where $E[\cdot]$ denotes the expectation operator and $p_{\mathbf{X}}(\mathbf{X})$ denotes the joint probability distribution of N numbers of input variable ($\mathbf{X}$). To compute Equation (4-1), an expensive multivariate integration ($I[\cdot]$) is required, such as

$$I[y^m(\mathbf{X})] \equiv \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \{y(x_1, \dots, x_N\}^m dx_1 \cdots dx_N \tag{4-2}$$

To ease the expensive computational cost of the multivariate integration, approximate integration methods have been developed; the univariate dimension reduction (UDR) method provides a representative outcome. The UDR method utilizes a dimension reduction method using additive decomposition. The additive decomposition converts a multivariate equation ($y(\mathbf{X})$) into a univariate approximation (y_{a}), given as

$$y(x_1, \dots, x_N) \cong y_{\text{a}}(x_1, \dots, x_N) = \sum_{j=1}^{N} y\big(\mu_1, \dots, \mu_{j-1}, x_j, \mu_{j+1}, \dots, \mu_N\big) - (N-1) y(\mu_1, \dots, \mu_N) \tag{4-3}$$

where μ_j is the first statistical moment of μ_j, $y(\mu_1, \dots, \mu_{j-1}, x_j, \mu_{j+1}, \dots, \mu_N)$ denotes a random system response that depends on x_j, and $y(\mu_j, \dots, \mu_j)$ is deterministic system response when $\mathbf{X} = \boldsymbol{\mu}$. Using the univariate approximation, multivariate integration of Equation (4-2) can be reduced to a univariate one, given as

$$I[y^m(\mathbf{X})] \cong \left[\sum_{j=1}^{N} I\left[y\left(\mu_1, \dots, \mu_{j-1}, x_j, \mu_{j+1}, \dots, \mu_N\right)\right] - (N-1)I[y(\mu_1, \dots, \mu_N)]\right]^m \tag{4-4}$$

Then, the m^{th} statistical moment of the predicted system response (y_{pre}) is approximately calculated as

$$\begin{aligned} E[y^m(\mathbf{X})] &\cong E[y_a^m(\mathbf{X})] \\ &= E\left[\left\{\sum_{j=1}^{N} y\left(\mu_1, \dots, \mu_{j-1}, x_j, \mu_{j+1}, \dots, \mu_N\right) - (N-1)y(\mu_1, \dots, \mu_N)\right\}^m\right] \\ &= \int_{-\infty}^{+\infty} \left\{\sum_{j=1}^{N} y\left(\mu_1, \dots, \mu_{j-1}, x_j, \mu_{j+1}, \dots, \mu_N\right) - (N-1)y(\mu_1, \dots, \mu_N)\right\}^m \\ &\qquad \cdot p_{x_j}(x_j)dx_j \end{aligned} \tag{4-5}$$

where one multivariate integration is converted to multiple univariate integrations. To calculate Equation (4-5), a binomial formulate can be used to recursively execute univariate integration, which requires using $(M - 1) \times N + 1$ integration points, where N is the number of input variables and M is the number of integration points along each random variable (Rahman and Xu 2004). To carry out each univariate integration effectively, the eigenvector dimension reduction method proposes $2M + 1$ and $4M + 1$ eigenvector sampling schemes, while maintaining high accuracy. For further information about univariate dimension reduction and eigenvector dimension reduction, please refer to (Rahman and Xu 2004) (Youn et al. 2008) (Youn and Wang 2008).

The computational cost for the UP can be decreased by using dimension reduction methods, however, the accuracy issue remains. According to (Youn et al. 2008), the numerical error in the use of the additive decomposition (Equation (4-4)) is given as

$$T\big[E[y^m(\mathbf{X})]\big] - T\big[E[y_\mathrm{a}^m(\mathbf{X})]\big] = \frac{1}{2!\,2!}\sum_{i<j} \frac{\partial^4 y^m}{\partial x_i^2 \partial x_j^2}\bigg|_{x_i, x_j = 0} E[x_i^2 x_j^2] + \cdots \tag{4-6}$$

where $T[\cdot]$ denotes the multivariate Taylor expansion at $\mathbf{X} = \mathbf{0}$. For example, when propagating the uncertainty in two input variables (x_1 and x_2) of a computational model ($y = x_1 x_2$), the numerical error exists in obtained statistical moments, as summarized in Table 4-3.

Table 4-3 Numerical Errors for Using Dimension Reduction in Calculation of the First Four Statistical Moments

The Statistical Moments	**Numerical Errors for $y = x_1x_2$**
1st order (m = 1)	$= \frac{1}{2!\,2!}\sum_{i<j}\frac{\partial^4 x_1x_2}{\partial x_1^2\partial x_2^2}(\mathbf{0})\mathrm{E}(x_1x_2) = 0$
2nd order (m = 2)	$= \frac{1}{2!\,2!}\sum_{i<j}\frac{\partial^4 (x_1x_2)^2}{\partial x_1^2\partial x_2^2}(\mathbf{0})\mathrm{E}((x_1x_2)^2) + \cdots = \frac{1}{4}\frac{\partial^4 (x_1^2x_2^2)}{\partial x_1^2\partial x_2^2}\mathrm{E}(x_1^2x_2^2)$ $= \frac{1}{4}(4)\mathrm{E}(x_1^2x_2^2) = \mathrm{E}(x_1^2x_2^2)$
3rd order (m = 3)	$= \frac{1}{2!\,2!}\sum_{i<j}\frac{\partial^4 (x_1x_2)^3}{\partial x_1^2\partial x_2^2}(\mathbf{0})\mathrm{E}((x_1x_2)^3) + \frac{1}{3!\,3!}\sum_{i<j}\frac{\partial^6 (x_1^3x_2^3)}{\partial x_1^3\partial x_2^3}\mathrm{E}(x_1^3x_2^3) + \cdots$ $= \frac{1}{36}\frac{\partial^4 (x_1^3x_2^3)}{\partial x_1^3\partial x_2^3}\mathrm{E}(x_1^3x_2^3) = \frac{1}{36}(36)\mathrm{E}(x_1^3x_2^3) = \mathrm{E}(x_1^3x_2^3)$
4th order (m = 4)	$= \frac{1}{4!\,4!}\sum_{i<j}\frac{\partial^4 (x_1x_2)^4}{\partial x_1^4\partial x_2^4}(\mathbf{0})\mathrm{E}((x_1x_2)^4) = \frac{1}{4!\,4!}\frac{\partial^4 (x_1^4x_2^4)}{\partial x_1^2\partial x_2^2}\mathrm{E}(x_1^4x_2^4)$ $= \frac{1}{4!\,4!}(4!\,4!)\mathrm{E}(x_1^4x_2^4) = \mathrm{E}(x_1^4x_2^4)$

4.3.3 Calibration Metrics for SOUP

The proposed SOUP method uses two UP methods in sequential formulation of two single loops. As explained in Section 3, the results of each UP method are different. The result from direct MCS is a probability distribution; the result from UDR is a set of statistical moments. Thus, two different calibration metrics that quantify the agreement or disagreement between experimental observations and computational predictions should be devised in two sequences. This study adopts 1) the moment matching metric (MMM) for the 1st sequence of SOUP and 2) the probability residual (PR) for the 2nd sequence of SOUP.

4.3.3.1. The Moment Matching Metric

The Moment Matching Metric (MMM) quantifies the difference between the statistical moments of the system responses from the observations and predictions. Thus, the MMM can be used with the UDR, of which the UP results are in statistical moments. The MMM, formulated either by the absolute difference or the squared difference, is defined as

$$MMM_{\text{abs}} = \sum_{i=1}^{4} \left| \boldsymbol{\theta}_{\mathbf{Y}_{\text{obs}},\, i} - \boldsymbol{\theta}_{\mathbf{Y}_{\text{pre}},\, i} \right| \tag{4-7}$$

$$MMM_{\text{squ}} = \sum_{i=1}^{4} \left(\boldsymbol{\theta}_{\mathbf{Y}_{\text{obs}},\, i} - \boldsymbol{\theta}_{\mathbf{Y}_{\text{pre}},\, i} \right)^2 \tag{4-8}$$

where $\boldsymbol{\theta}_{\mathrm{Y}}$ denotes the statistical moments of the system responses (Y) from the observations and predictions. Both the absolute and the squared difference-based MMM calculate the differences in the first four statistical moments (mean, standard deviation, skewness, kurtosis). In general, the first four statistical moments are enough to describe the uncertainty or variability in the outputs. For example, the

Pearson system is available for characterizing a probability distribution using the first four statistical moments (Youn and Wang 2008) (Youn and Xi 2009) (Choi et al. 2010). However, the first four statistical moments are insufficient information for a full comparison of the probability distributions. For a perfect match between the observations and predictions, the MMM gives a 'zero' value, which means no difference.

4.3.3.2. The Probability Residual

The PR quantifies the difference between the two probability distributions that are derived from the observations and predictions (Lee et al. 2018). The PR formulated either by the integral of the absolute difference or squared difference is defined as

$$PR_{\text{abs}} = \int_{-\infty}^{\infty} |p_{\mathbf{Y}_{\text{obs}}}(y) - p_{\mathbf{Y}_{\text{pre}}}(y|\boldsymbol{\theta}_{\mathbf{X}})|dy \tag{4-9}$$

$$PR_{\text{squ}} = \int_{-\infty}^{\infty} \left(p_{\mathbf{Y}_{\text{obs}}}(y) - p_{\mathbf{Y}_{\text{pre}}}(y|\boldsymbol{\theta}_{\mathbf{X}})\right)^2 dy \tag{4-10}$$

where $p_{\mathbf{Y}\text{obs}}(y)$ denotes the PDF of the system response from the experimental observations and $p_{\mathbf{Y}\text{pre}}(y|\boldsymbol{\theta})$ denotes the PDF of the system responses from the computational predictions. For a perfect match between the observations and predictions, the PR also gives a 'zero' value, which means no difference. Unlike the MMM, the PR can expect accurate model calibration, in that it can quantify the overall shape difference of the PDF. For accurate OBMC, accurate one-to-one comparison between two probability distributions from observations and predictions is desired. In this context, the PR is a better choice than the MMM.

4.3.3.3. Summary and Discussion

One desired characteristic of a calibration metric is whether it can formulate a convex optimization problem. When an optimization problem is confirmed to be convex, it can be solved accurately and efficiently using existing optimization tools (Boyd and Vandenberghe 2004). However, the previous study by Lee et al. (Lee et al. 2018) proved that OBMC with PR cannot formulate a convex optimization problem. On the other hand, the MMM can formulate a convex optimization for OBMC. For the MMM, the determinant of the Hessian matrix is always positive. The convexity of OBMC using the MMM allows the optimization to always converge to the optimum.

4.4 Summary and Discussion

In conclusion, SOUP can grasp both the efficiency and accuracy of available methods by using two UP methods with two calibration metrics. The first sequence in SOUP uses UDR to obtain the statistical moments of the predicted system responses and compares the observations and predictions using the MMM. Due to the efficiency of UDR, the computational time for the first sequence is fast. By using the MMM, it can be expected that the optimization converges at the end of the first sequence. While numerical errors from the UDR and incomplete comparison from the MMM may cause inaccurate OBMC in the first sequence, these issues can be improved in the second sequence of SOUP. The global optimum at the end of the first sequence becomes the intermediate calibration parameter, the initial calibration parameter for the second SOUP sequence. Although the computational cost of direct MCS is high, it can be confirmed that the intermediate calibration parameter from the second sequence with UDR and the MMM reaches the near-true optimum. In other words, the number of iterations required for the second sequence and the optimization can be conducted in the convex zone. Since the second sequence with direct MCS and the PR shows the best accuracy, accurate calibration can be expected at the end of the two-step SOUP process.

Sections of this chapter have been published or submitted as the following journal article:

1) **Guesuk Lee** and Byeng D. Youn, "Sequential Optimization and Uncertainty Propagation Method for Efficient Optimization-Based Model Calibration," *Structural and Multidisciplinary Optimization*, Submitted in September 2018.

Chapter 5

Development of a Statistical Validation Metric –Probability of Coincidence

Model validation is the process of determining the degree to which a computational model is an accurate representation of a real phenomenon, from the perspective of the model's intended uses (Babuska and Oden 2004) (Oberkampf et al. 2004b) (Trucano et al. 2006) (Hills et al. 2008) (Oberkampf and Trucano 2008) (Kutluay and Winner 2014). In this dissertation, model validation is implemented to check whether calibration has been conducted accurately at the completion of model calibration. (Note that validation experiments are designed to validate the calibrated results with different types of system responses.) However, model validation can be a stand-alone process to assess the credibility of a computational model with or without the calibration process. In addition, the validation results can help the process of improving the model. Like statistical model calibration, the process of model validation should also be conducted in a statistical sense. The objective of Research Thrust 4 is thus to develop a way to statistically validate the calibrated computational model.

First, Research Thrust 4 begins with a review of studies of the developed validation methods. In particular, the review focuses on the area metric with hypothesis testing. Due to its several merits, many prior studies have adopted the area metric with the hypothesis testing for statistical model validation. First, the area metric quantifies the statistical difference between computational predictions and experimental observations. Second, the hypothesis testing decides whether or not the prediction of the computational model is acceptable based on the evaluation of the area metric. Nonetheless, several limitations of the area metric with hypothesis testing remain; these remaining issues motivated Research Thrust 4. Research Thrust 4 thus elects an appropriate validation metric for validating the calibrated results by OBMC. At the end, a new validation metric is proposed, which is motivated by the previous studies.

The remainder of Chapter 5 is organized as follows. Section 5.1 reviews statistical model validation methods developed from the field of verification and calibration (V&V) (Oberkampf and Roy 2010). In Section 5.2, a validation metric – probability of coincidence (POC) – is proposed for the validity check of the calibrated results derived from OBMC. Finally, the conclusions of Research Thrust 4 are provided in Section 5.3.

5.1 Brief Review of Statistical Model Validation

This section provides a review of statistical model validation. Earlier studies by Oberkampf and Barone (Oberkampf and Barone 2006), Ferson et al. (Ferson et al.

2008), and Liu et al. (Liu et al. 2011) summarized the desired features of validation metrics. Based on these studies, the desired features of a validation metric can be summarized as "objective" or "stochastic or statistical." Furthermore, in a statistical sense, a system response is presented as a variation, for example, by a probability distribution or a random process, due to the existence of various uncertainties. A "statistical or stochastic" validation metric should be able to compare the system responses from experiments and simulations in a variation by considering uncertainties (Schwer 2007) (Sarin et al. 2010) (Jiang and Mahadevan 2011). For statistical comparison, the stationary type of system responses needs a distribution comparison method because the scalar values build a distribution from different conditions of prediction and observation (Chen et al. 2004) (Mahadevan and Rebba 2005) (Halder and Bhattacharya 2011). This dissertation focuses on the model validation method for the stationary type of system responses; however, there also exists a need for model validation methods for the dynamic type of system responses.

Normally, the developed methods of statistical model validation are composed of the validation metric (Section 5.1.1) and the decision problem (Section 5.1.2). At the end of a review of each of these, Section 5.1.3 summarizes the conclusions.

5.1.1 Validation Metric – The Area Metric

As a powerful validation metric, numerous studies in the model validation community show that the area metric possesses most of the desirable features of a validation metric for comparing experimental observations and computational predictions (Ferson et al. 2008) (Ferson and Oberkampf 2009) (Liu et al. 2011) (Roy and Oberkampf 2011) (Thacker and Paez 2013) (Voyles and Roy 2015). To obtain the area metric value, propagated system responses of computational prediction produce a probability box or p-box ($F(y)$), while n numbers of experimental measurements are used to construct an empirical cumulative density function (CDF, $S_n(y)$) of system responses (Ferson et al. 2008). Finally, the minimum area between these two structures is referred to as the value of the area metric (d_{area}), shown as:

$$d_{\text{area}}\big(F(y), S_n(y)\big) = \int_{-\infty}^{\infty} |F(y) - S_n(y)|dy \tag{5-1}$$

Prior comparative studies (Oberkampf and Barone 2006) (Ferson et al. 2008) (Liu et al. 2011) (Roy and Oberkampf 2011) (Thacker and Paez 2013) (Li et al. 2014), showed the area metric to be promising due to its favorable features, as compared to other methods. First, the area metric is one of only a few developed distribution comparison metrics. It measures the entire distribution, rather than statistical moments, thereby accounting for uncertainties in both the simulation and the experiments. Second, sampling uncertainty that arises due to limited experimental data available for the validity check is considered in the area metric (Jung et al. 2014).

U-pooling and T-pooling methods are proposed to assist with usage of the area metric (Ferson et al. 2008) (Liu et al. 2011) (Li et al. 2014). While the area metric can validate the prediction with a small number of data, conducting a few more experiments could still be a burden to the analysts. The U-pooling method can ease this burden by pooling all experimental observations at different validation sites into a u-value CDF (Ferson et al. 2008) (Liu et al. 2011) (Li et al. 2014). Through the u-pooling technique, the area metric takes advantage when multiple experiments at various validation sites are available. The original u-pooling method is only applicable for a single-system response at a single validation site. To extend the usage of the area metric for validating correlated multiple responses, Li et al. (Li et al. 2014) proposed the t-pooling method, accompanied by probability integral transformation. The multivariate probability integral transformation method transforms the joint CDF of the system responses into a univariate CDF; then, the t-pooling method integrates the evidence from all relevant data of multi-response quantities over an intended validation domain into a single measure to assess the overall disagreement. The area metric with U-pooling and T-pooling methods is attractive in that it uses all possible data to validate a model. However, defining the range of validation sites where the data are pooled in one distribution is an important challenge.

5.1.2 Decision Problem - Hypothesis Testing

A validation metric is a stand-alone measure that indicates the degree of agreement or disagreement between computational predictions and experimental observations. (Note that the area metric quantifies the minimum disagreement between computational predictions and experimental observations.) However, the acceptance criteria for a yet-to-be-validated model is another important issue. Suppose a desirable validation metric quantified the difference between the experiment and the simulation. It is then necessary to establish a method to determine the validity of the model (Oberkampf and Trucano 2002).

Statistical hypothesis testing with a specified level of significance is widely used as a decision-making tool for model validation (Chen et al. 2004) (Liu et al. 2011) (Kokkolaras et al. 2013) (Jung et al. 2014). Hypothesis testing aims to determine whether the acceptance or rejection of a model is valid or not using quantitative measurements of the discrepancy between the experiment and the simulation. Two kinds of hypothesis testing have been studied: classical and Bayesian (Oberkampf and Barone 2006) (Jiang and Mahadevan 2007) (Jiang and Mahadevan 2008) (Oliver et al. 2015) (Li and Mahadevan 2016). The significant difference between the two testing methods is that classical testing focuses on model rejection in the validity check; whereas, the Bayesian method focuses on model acceptance by using prior information (Liu et al. 2011). This dissertation, however, focuses on discussing the classical hypothesis testing.

Classical hypothesis testing is a well-developed statistical method for accepting or rejecting a model's validity based on statistics (Chen et al. 2004) (Oberkampf and Barone 2006) (Kat and Els 2012) (Ling and Mahadevan 2013) (Jung et al. 2014). Table 5-1 summarizes the hypothesis testing. First, the null hypothesis (H_0, which represents that a computational model is valid) and the alternative hypothesis (H_1, which represents that a computational model is not valid) are defined. The former means that the difference between the predicted and observed system responses is not statistically significant; the latter means there is a statistically significant difference. Hypothesis testing is based on test statistics ($\boldsymbol{\theta}$). If the test statistic falls outside of the critical region of the test statistic, the null hypothesis is rejected, meaning that the observations from experiments and the simulation prediction are significantly different. To use the area metric, the calculated area metric becomes a test statistic. Figure 5-1 shows an example of model validation with the area metric (d_{area}) and hypothesis testing. The corresponding P-value ($P_{\text{area}}(0.05) = 0.137$, when the number ($n$) of experimental observations is 18 and the significance level is 0.05, in Figure 5-1) is calculated as the probability that the test statistic will fall outside the range defined by the calculated value of the test statistic under the null hypothesis

(Rebba and Mahadevan 2006) (Ling and Mahadevan 2013).

According to Liu et al. (Liu et al. 2011), classical hypothesis testing seldom rejects a better model. A problem arises when a small amount of physical experiment data is available; this is common in many fields. Comparing full distribution data is impossible, but with a small number of data there is possibility that the result is not trustworthy. In this situation, the confidence interval of prediction distribution can be used by checking if the observed data fall inside the interval (Halder and Bhattacharya 2011) (Ghanem et al. 2008) (Buranathiti et al. 2006) (Chen et al. 2004). However, this strategy can tend not to reject an incorrect model, since a small number of data fall inside the confidence interval of prediction with high possibility (Liu et al. 2011). Furthermore, specifying the confidence level creates another problem, since a small perturbation of the confidence level largely affects the results of acceptance or rejection. On the other hand, a large number of samples can give

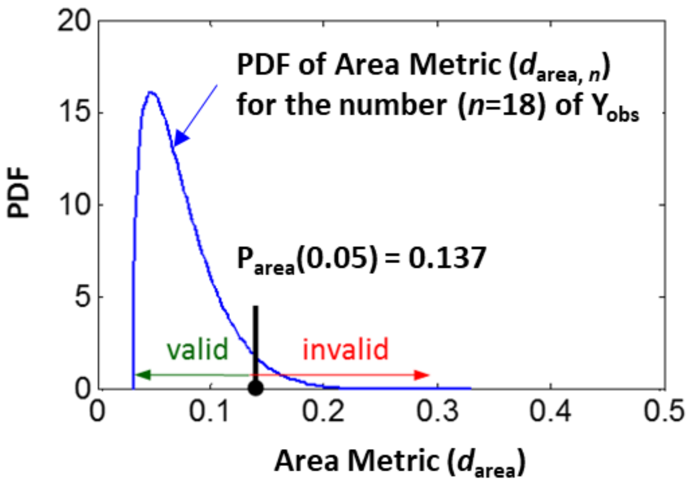

Figure 5-1 An Example of a Probability Density Function of the Area Metric's Value for n =18 Number of Experimental Observations

misleading results, because as the number of samples increases, the null hypothesis tends to be rejected at a given significance level (Ling and Mahadevan 2013). Above all, it should be noted that failing to reject the null hypothesis does not prove that the null hypothesis is true.

Table 5-1 Summary of Hypothesis Testing

The Null Hypothesis (H_0)	A computational model is valid. (e.g., The area metric value – the difference between experimental observations and computational predictions – is significantly small.)
The Alternative Hypothesis (H_1/H_a)	A computational model is not valid. (e.g., The area metric value – the difference between experimental observations and computational predictions – is significantly large.)
False Positive (Type 1 Error)	Incorrect rejection of a true null hypothesis (e.g., a computational model is judged as not valid, although it is truly valid.)
False Negative (Type 2 Error)	Failure to reject a false null hypothesis (e.g., a computational model is judged as valid, although it is truly not valid.)
True Positive	Successfully not rejecting a true null hypothesis
True Negative	Successfully rejecting a false null hypothesis
Significance Level (P-value/Probability of Type 1 Error)	The probability of a value accurately rejecting the null hypothesis

5.1.3 Summary and Conclusion

So far, a comprehensive review of the area metric and hypothesis testing, respectively, as a validation metric and decision making tool, was introduced. Despite several merits of using the area metric with hypothesis testing, there are two limitations of its use as a statistical validation method for OBMC.

First, the area metric cannot quantify the difference between PDFs from experimental observations and the computational predictions that are characterized by Research Thrust 1. In a probabilistic context, it is generally necessary to quantify the variability in QOI due to various uncertainty factors using a probability distribution. In the first part of OBMC (Research Thrust 1), the uncertainty characterization of experimental observations results in a parametric PDF. Characterized variability derived from experimental observations can be applied to both calibration and validation processes. For calibration by OBMC, the optimization progresses toward maximizing the agreement between the PDFs of experimental observations and computational predictions. Also, for validation, the uncertainty characterization resulting from the proposed method in Research Thrust 1 can be used. With the proposed uncertainty characterization method that considers different experimental conditions as systematic measurement errors, experimental observations from different experiment settings are integrated into one parametric PDF. Thereby, the statistical uncertainty that arises from a dearth of data can be eased. However, to use the results of the method proposed in Research Thrust 1, a validation metric should be able to quantify the difference between the PDFs derived from the experimental observations and computational predictions. However, the area metric is computed by the minimum area between a probability box or p-box that comes from the propagated system responses of prediction and an empirical CDF derived from the experimental observations.

Second, the decision by hypothesis testing based on calculated values of the area metric does not provide the justification that a computational model is valid. Four decisions can be made based on the hypothesis testing: 1) false positive (type 1 error); 2) false negative (type 2 error); 3) true negative (correct decision), or; 4) true positive (correct decision). First, wrong decisions, or statistical errors, can be made in statistical hypothesis testing, as summarized in Figure 5-2. A *false positive* (type 1 error) denotes an incorrect rejection of a true null hypothesis (H_0) that a computational model is valid, which means that the system response from a computational model follows the system response from the experiments; an alternative hypothesis (H_1) stands for the opposite. A *false negative* (type 2 error) is the failure to reject a false null hypothesis. A wrong decision, or a statistical error (type 1 and 2 error) can be adjusted by the criteria, the rate of type 1 or 2 error. For example, as type 1 error (or significance level, α) increases, type 2 error decreases. By adjusting a type 1 error to be high (high significance level), the probability of accurately rejecting the null hypothesis becomes high, which means the decision can be made conservatively. Likewise, there are two possible correct decisions. A *true positive* decision denotes success in not rejecting a true null hypothesis. A *true negative* decision denotes success in rejecting a false null hypothesis. The decision of *true negative* states that the computational model is not valid for its intended use. However, the decision of *true positive* does not strictly state that the computational is definitely valid for its intended use. This is because the hypothesis testing is formulated only to reject an invalid computational model.

In summary, the area metric with hypothesis testing has limitations when it is used in OBMC, as advanced in this dissertation, specifically: 1) the area metric

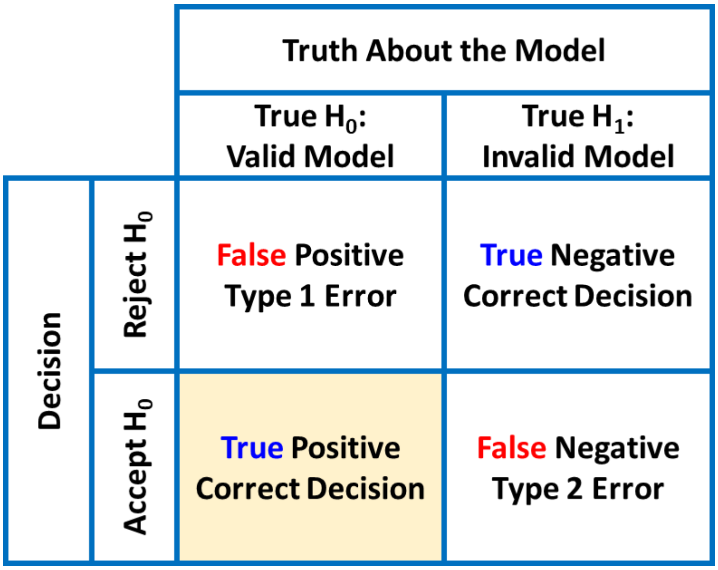

Figure 5-2 Decisions by Hypothesis Testing: 1) False Positive (Type 1 Error); 2) False Negative (Type 2 Error); 3) True Negative (Correct Decision); 4) True Positive (Correct Decision)

cannot be used to quantify the difference between PDFs from experimental observations and computational predictions, and; 2) a true positive decision that results from the hypothesis testing does not prove a computational model is valid for its intended use.

5.2 Probability of Coincidence

This section introduces the proposed statistical validation method – probability of coincidence (POC). First, Section 5.2.1 explains the motivations behind the proposed POC. Second, Section 5.2.2 presents the mathematical formulation of the proposed POC.

5.2.1 Motivations Behind the Probability of Coincidence Method

Probability of coincidence (POC) is a new validation metric motivated by the idea of 1) probability of failure for reliability analysis (Ichikawa 1993) (Hu et al. 2018) and 2) probability of separation (POS) for a class separability metric (Jeon 2016). Taking into account that failure occurs when an applied load exceeds the strength of a system, the probability of failure (p_{f}) can be described as the probability that the load or stress (L) exceeds the strength (S), shown as

$$p_{\mathrm{f}}(x) = \int_{0}^{\infty} F_S(x) f_L(x) dx \tag{5-2}$$

where F_S and f_L denote the CDF of strength (S) and the PDF of stress or loading (L), respectively. Figure 5-3 depicts the probability of failure, which is calculated by multiplying the probability ($f_L(x^*)$) that a loading has x^* value and the probability ($F_S(x^*)$) that the strength is lower than the x^* value.

The POS is also motivated by the probability of failure. First, the probability (P_{NS}) of a non-separable region is computed as

$$P_{\mathrm{NS}} = \int_{-\infty}^{\infty} F_{\mathrm{c2}}(x) f_{\mathrm{c1}}(x) dx \quad for \;\; \tilde{x}_{\mathrm{c1}} \leq \tilde{x}_{\mathrm{c2}} \tag{5-3}$$

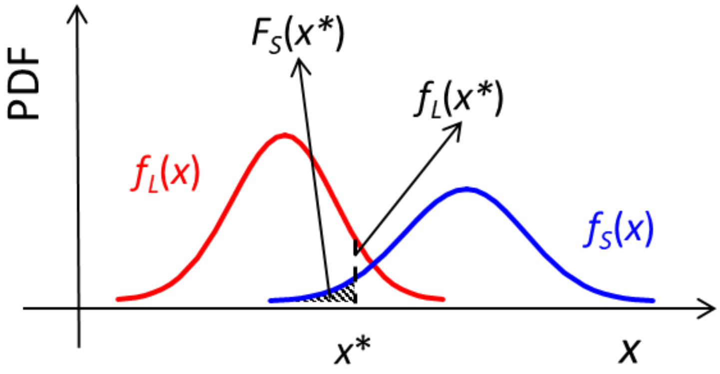

Figure 5-3 Illustration of the Failure by Probability Density Functions of Loading and Strength

where f_{c1} and F_{c2} represent the PDF of class 1 and the CDF of class 2, respectively, while $\tilde{x}_{c1}$ and $\tilde{x}_{c2}$ correspond to the medians of classes 1 and 2, respectively. P_{NS} should have a value range between 0 and 0.5. Thus, to have a normalized value between 0 and 1, the POS is formulated as

$$POS = (e^{(1-2P_{\mathrm{NS}})} - 1)/(e - 1) \tag{5-4}$$

Finally, the POS can have a "0" value if the feature data of two different classes overlaps perfectly and can have a "1" value if not overlapped at all.

5.2.2 Formulation of Probability of Coincidence

Motivated by the idea that the POS can have a “0” value if the feature data of two different classes overlaps perfectly and a “1” value if not overlapped at all, the probability of coincidence (POC) is presented as

$$POC = \left(1 - \frac{e^{(1-2P_{\mathrm{ov}})} - 1}{e - 1}\right) \times 100 \ \ [\%] \tag{5-5}$$

where the probability of the overlapped area (p_{ov}) is presented as:

$$\begin{gathered} P_{\mathrm{ov}} = \int_{-\infty}^{\infty} P_{\mathrm{obs}}(y) p_{\mathrm{pre}}(y) dy \\ = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{y} p_{\mathrm{obs}}(y)\, dy \right) p_{\mathrm{pre}}(y) dy \\ for \ \ \tilde{y}_{\mathrm{obs}} \leq \tilde{y}_{\mathrm{pre}} \end{gathered} \tag{5-6}$$

$$\begin{gathered} P_{\mathrm{ov}} = \int_{-\infty}^{\infty} P_{\mathrm{pre}}(y) p_{\mathrm{obs}}(y) dy \\ = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{y} p_{\mathrm{pre}}(y)\, dy \right) p_{\mathrm{obs}}(y) dy \\ for \ \ \tilde{y}_{\mathrm{pre}} \leq \tilde{y}_{\mathrm{obs}} \end{gathered} \tag{5-7}$$

Figure 5-4 describes the probability of the overlapped area (p_{ov}). Equation (5-6) can be described by Figure 5-4, in that p_{ov} is calculated by the integral of multiplication of the shaded area ($Py_{\mathrm{pre}}(y^*)$) and the dotted length ($py_{\mathrm{pre}}(y^*)$). Finally, the POC can have a 0 [%] percentage value if the two probability distributions from the observations and predictions show no match; and 100 [%] percentage value if the two probability distributions show a perfect match. The validation results from the POC can provide valuable information to the analysts who make the decision whether to believe the credibility of a computational model about the probability (percentage) of the agreement between the observations and predictions.

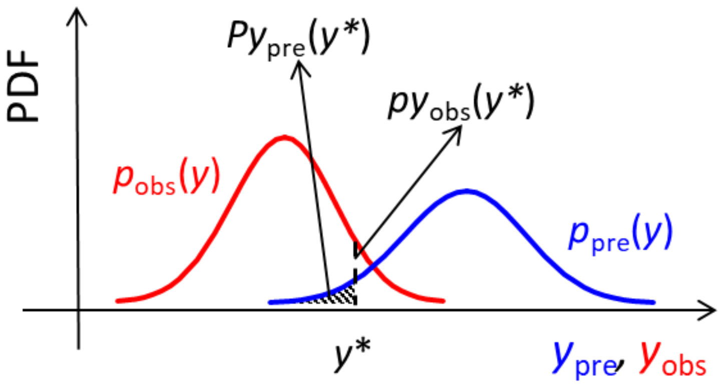

Figure 5-4 Illustration of the Probability of Overlapped Area Derived from Probability Density Functions of Computational Predictions and Experimental Observations

5.3 Summary and Discussion

Research Thrust 4 aims to develop a method for statistical model validation, which is the final process of OBMC. Statistical model validation checks the calibrated results using a validation metric and a decision method. Chapter 5 begins by providing a comprehensive review of the limitations of the area metric and hypothesis testing. Research Thrust 4 attempts to develop a new validation method for alternating the area metric and hypothesis testing. Ultimately, a new validation metric POC is proposed. POC itself quantifies the agreement and provides probabilistic information to the analysts who make a decision about whether or not to use a computational model. However, there is still a room for developing other validation methods. Suggestions for future study are discussed in the conclusion section of this dissertation.

Sections of this chapter have been published or submitted as the following journal article:

1) Guesuk Lee, Wongon Kim, Hyunseok Oh, and Byeng D. Youn, Nam H. Kim "Review of Statistical Model Calibration and Validation – From the Perspective of Uncertainty Structures," *Structural and Multidisciplinary Optimization*, Submitted in September 2018.

Chapter 6

Case Studies

To demonstrate the four proposed research thrusts, this chapter employs four case studies, as summarized in Table 6-1, specifically: 1) uncertainty characterization of experimental observations of a cantilever beam deflection (Section 6.1), 2) model calibration of a bearing capacity equation (Section 6.2), 3) model calibration and validation of a steering wheel's vibrational model (Section 6.4), and 4) model calibration and validation of a thin-film transistor liquid crystal display's deflection model (Section 6.3). The first two case studies are introduced to demonstrate the effectiveness and accuracy of the proposed methods using problems with known answers. The final two cases studies are implemented to check the applicability of the proposed methods using actual, real-world engineering examples.

Table 6-1 Summary of Case Studies

	Research Thrust 1	Research Thrust 2	Research Thrust 3	Research Thrust 4
Case Study 1: Cantilever Beam's Deflection	∨ Check the accuracy with a known answer	-	-	-
Case Study 2: Bearing Capacity Equation	-	∨ Check the accuracy with a known answer	∨ Check the accuracy with a known answer	-
Case Study 3: Steering Wheel Vibrational Model	∨ Check the applicability in a practical example	-	∨ Check the applicability in a practical example	∨ Check the applicability in a practical example
Case Study 4: Liquid Crystal Display Deflection Model	∨ Check the applicability in a practical example	-	∨ Check the applicability in a practical example	∨ Check the applicability in a practical example

6.1 Case Study 1: Uncertainty Characterization of Observed Cantilever Beam Deflection

As a case study, a cantilever beam problem is adopted to check the accuracy of Research Thrust 1. Section 6.1.1 presents the problem description of Case Study 1. Section 6.1.2 presents the estimated results of the probability density functions for the variability and two measurement errors in the observed deflections. For data of the beam's deflection, the author intentionally generated sample data. At the end, Section 6.1.3 summarizes the results of Case Study 1.

6.1.1 Problem Description

The subject of the case study is a cantilever beam horizontally protruding from a rigid, unyielding vertical wall (Figure 6-1). As a QOI, the problem seeks to define the cantilever beam's deflection (y_{D}) at the free end of the beam when a vertical downward loading (P_{D}) is applied at the same point. For repeated measurements, samples are randomly selected from a population of rectangular cantilever beams, of which the geometric (length, height, and width of a cantilever beam) and material properties (density and Young's modulus of a cantilever beam) vary among the population. In the problem, it is tentatively assumed that the beam is made of a homogeneous isotropic material, which denotes that the same value of the material properties can be attributed to the beam's deflection through the entire beam. (Note that there is no dependency or statistical correlation among the geometric and material properties.) Due to somewhat large variations in the geometric and material properties, the randomly selected beam shows the variability in the observed system response ($y_{\mathrm{D,\,obs}}$).

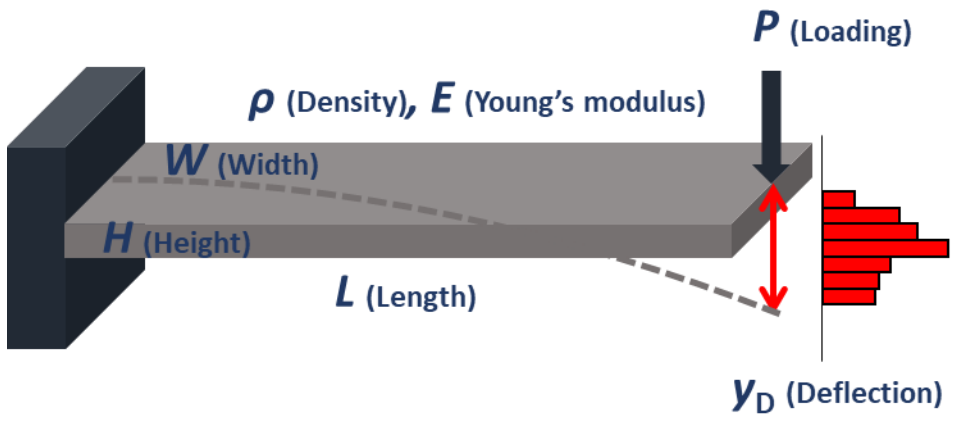

Figure 6-1 Illustration of Deflection of a Cantilever Beam When a Vertical Downward Loading is Applied at the Tip End of the Beam.

In the problem, the observed deflection ($y_{D,\ obs}$) is further subject to two measurement errors that arise during the pick-up and usage of a measuring sensor. Here, a measuring sensor is randomly picked from a population of sensors and the selected sensor is repeatedly used for a set of data. When the chosen sensor has biased accuracy, described by a distribution or range of error, then all the measurements of each set produced from that sensor may have similar biased error; this is called systematic measurement error. Meanwhile, random measurement errors may exist both inside the set of data and across the sets of data. For example, detachment of the chosen sensor for reuse in four different beams may cause random measurement errors through the entire data in a set. Also, the chosen sensor for each set of data can have random measurement errors, since the data are randomly picked from a sensor population, across which variability in accuracy may exist.

6.1.2 Uncertainty Characterization of Observed Deflection of a Cantilever Beam Considering Measurement Errors

The true mean and standard deviation of observed beam's deflections ($\mathbf{Y}_{\mathrm{D,\,obs}}$) are set to '0.1631' and '0.0153', respectively. (The unit for the deflection is omitted since it does not matter in developing the discussion.) To generate the observed data ($\mathbf{Y}_{\mathrm{D,\,obs}}$), including the two measurement errors, the maximum amount of bias 'α' for systematic measurement error and 'βy' standard deviation for random measurement are provided as '-0.1' (negative bias) and '$0.1y$', respectively (refer to Equations (2-3) and (2-4)). Second, increasing numbers ($n = \{3, 4, 5, 10, 25, 50\}$) of data are sampled to examine the statistical uncertainty that is present due to randomness in the generated random samples. For each number of sampled data, 30 iterative studies are conducted to check the accuracy. The method of moments, without considering the measurement errors in the observed data, is devised to check the effectiveness using the proposed method.

Figure 6-2 represents the estimated mean (μ_y) and standard deviation (σ_y) using repeatedly sampled (observed) deflection data. Each blue dot denotes an estimation with n numbers of sampled data. The upper parts (Figure 6-2a and Figure 6-2c) show the estimated results by the method of moments, without considering measurement errors; the lower figures (Figure 6-2b and Figure 6-2d) show the estimated results using the proposed method. In common for all figures in Figure 6-2, as the amount of observed data increases, the variations of the estimated results get narrower; this is because the larger amount of data alleviates the degree of statistical uncertainty that is present in earlier results due to a lack of data. By comparing Figure 6-2a and Figure 6-2b, it can be confirmed that the proposed method restores the biased results that are seen from the results found without the proposed method. Figure 6-2c shows that existence of random measurement error in observations results in larger estimations of standard deviation. On the other hand, Figure 6-2d shows that the proposed method adequately estimates the standard deviation by eliminating the effect of random measurement error.

6.1.3 Summary and Discussion

The method proposed in Research Thrust 1 is adopted for a case study of cantilever beam. Using the proposed method, more accurate estimations were available when both systematic and random measurement errors exist in the observations. Compared with a known answer for the estimations (the true mean and standard deviation of deflection of a cantilever beam), it could be confirmed that the estimated bias mean due to systematic measurement error was restored and the largely estimated standard deviation due to random measurement error was narrowed when the proposed method was used. With properly characterized variability of observed deflection available through use of the proposed method, the calibration will be more accurate. For example, if the biased estimation of the mean and largely estimated standard deviation were used for calibrating the unknown input variables, the calibrated input variables would have biased and largely calibrated results, which is untrue.

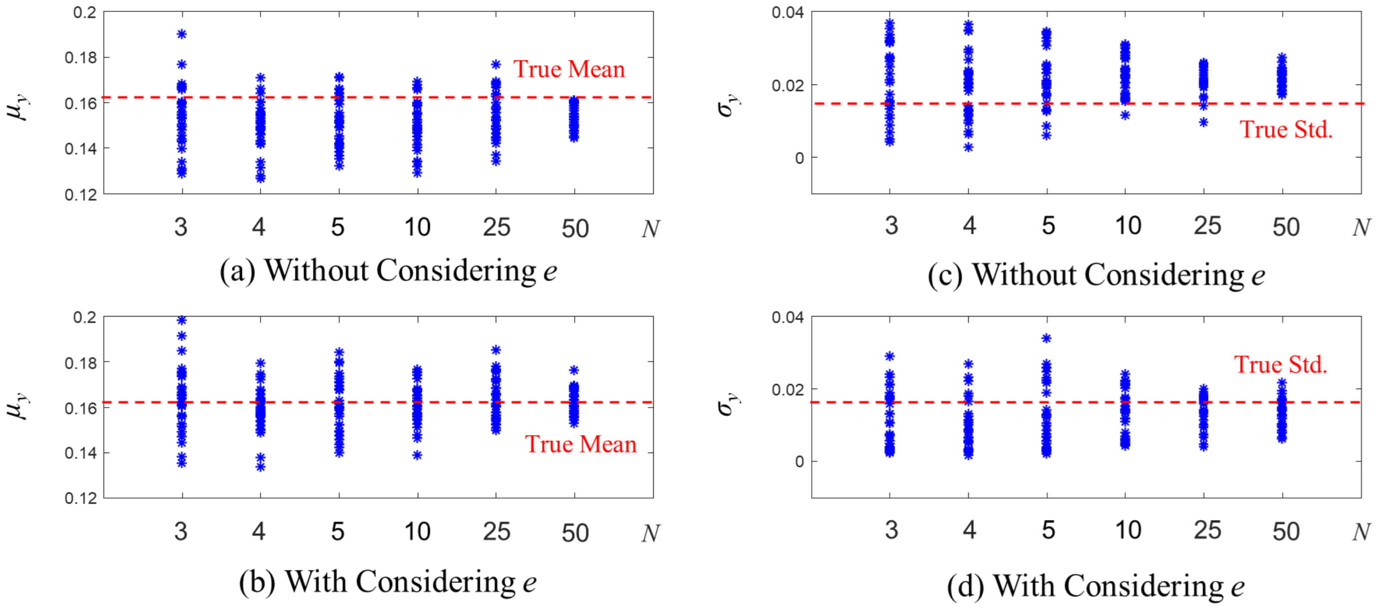

Figure 6-2 Estimated Mean and Standard Deviation with Increasing Numbers of Observations with/without Consideration of Measurement Errors

6.2 Case Study 2: Model Calibration of the Terzaghi's Bearing Capacity Equation

Section 6.2 introduces an example that examines a calibration problem of a bearing capacity equation for building a shallow foundation. The case study is formulated to check the effectiveness of the idea proposed in Research Thrusts 2 and 3. Section 6.2.1 presents the problem description of Case Study 2. Section 6.2.2 presents the calibrated results of the bearing capacity equation. First, the calibrated results check the effectiveness of the idea proposed in Research Thrust 2, OBMC with analytical sensitivity information. Next, the calibrated results check how the proposed SOUP (Research Thrust 3) improves the efficiency and accuracy of the calibration results by OBMC.

6.2.1 Problem Description

A textbook written by Halder and Mahadevan (Haldar and Mahadevan 2000) provides an example study for teaching quantification of uncertainty in the response when there are random input variables. A shallow foundation is a building technique that transfers building load to the earth near to the surface. A shallow strip footing, which is typical in ordinary buildings located in a dense sand layer, has a wider strip footing B_s than depth H_s from the ground surface (Figure 6-3). When designing a shallow strip footing to avoid general shear failure, the bearing capacity of the soil is an important system performance. General shear failure can result in sudden catastrophe associated with plastic flow and lateral expulsion of the soil. Terzaghi's bearing capacity (q_u) equation is a classic equation proposed by Terzaghi (Terzaghi 1944) that is still used in its original form, given as

$$q_u = c_s N_c + \gamma H N_q + 0.5\gamma B N_\gamma \quad (6\text{-}1)$$

where c_s is the cohesion, γ_s is the unit weight of the soil, and N_c, N_q, and N_γ are the bearing capacity factors that can be determined from the given information on the angle of internal friction of the soil.

This linear relationship between one response (q_u) and seven inputs (c_s, γ, N_c, N_q, N_γ, H, and B) is reformulated into a model calibration problem. In this model calibration problem, c_s, the soil cohesion, is assumed to be an unknown random input variable that follows a normal distribution, γ_s is a known random variable that follows a normal distribution, and the other inputs are constants. Assume that the soil layer has an angle of internal friction Φ of 20°, and the corresponding bearing capacity factors N_c, N_q, and N_γ are 17.7, 7.4, and 5.0, respectively. Further assume that γ_s = 1842 kg/m^3, H_s = 1.0 m, and B_s = 1.5 m. The properties of the seven input variables are shown in Table 6-2. The statistical information of experimental observations of the bearing capacity is given in Table 6-3. By using direct Monte

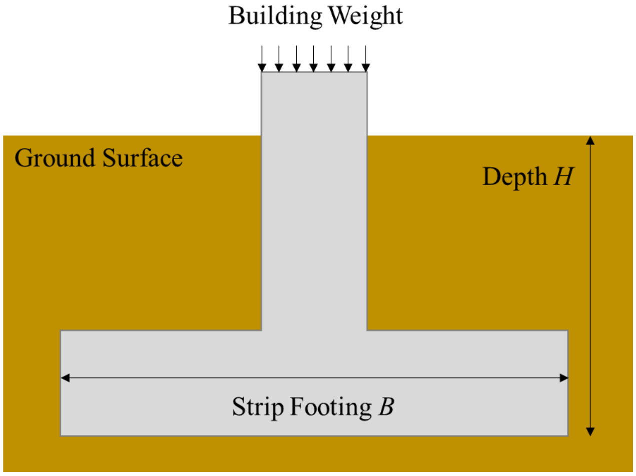

Figure 6-3 Scheme of a Shallow Strip Footing

Carlo Simulation method, the predicted PDF of the bearing capacity (q_u) can be obtained. For the unknown input variable (Table 6-2), the soil cohesion (c_s) and experimental observations (Table 6-3), bearing capacity (q_u), two cases are designed to check the effectiveness of the proposed method. The first case presents a smaller value of standard deviation for the unknown input variable and the second case presents a larger value of standard deviation for the unknown input variable.

Note that Terzaghi's bearing capacity equation is explicit; thus, the inverse problem of this equation can be solved analytically in a straightforward manner. In Table 6-2, values given for the cohesion in round brackets are the expected calibrated point (i.e., the exact solution); these values are calculated through the analytical function. The values can be compared with calibrated solutions computed using general methods, and with the proposed approaches for OBMC.

6.2.2 Model Calibration by Optimization-Based Model Calibration

Two subcases are formulated for Research Thrusts 2 and 3. The first subcase in Section 6.2.2.1 is formulated to compare the calibrated results from OBMC 1-1) without analytical sensitivity information and 1-2) with analytical sensitivity information (Research Thrust 2). Next, the second subcase in Section 6.2.2.2 is formulated to compare the calibrated results from 2-1) the basic single loop with MCS and 2-2) the proposed SOUP (Research Thrust 3). For the first subcase, the value of the unknown input variable and experimental observations are set to the smaller values shown in Table 6-2 and Table 6-3. For the second subcase, the value of the unknown input variable and experimental observations are set to both the smaller and larger values shown in Table 6-2 and Table 6-3.

6.2.2.1. Subcase 1: Optimization-Based Model Calibration with Analytical Sensitivity Information

The calibrated results from four different initial points are summarized in Table 6-4 and Table 6-5. The four initial points are selected based on the analytical information derived in Section 3.3.1, such as the inflection separating the convex and non-convex area of the response surface of the calibration metrics with respect to the calibration parameters. Initial point #1 lies on the convex of the likelihood function and PR. Initial point #2 and #3 lie on the convex of the likelihood function, but on the non-convex of the PR. Initial point #4 lies on the non-convex of the likelihood function and PR. The four initial points and global optimum are depicted in Figure 6-4.

Table 6-4 shows the calibrated results from OBMC without analytical sensitivity information, which means the finite difference gradient calculation is used to obtain the approximated sensitivity information. The red-shaded case denotes an initial point that lies on the non-convex area; the blue-shaded case denotes an initial point that lies on the convex area. Only when the OBMC from initial point #1 and the calibration metric is PR, do the calibrated results show robustness and accuracy. The results in Table 6-4 imply that the calibrated results may not be robust and accurate, even though OBMC begins from the convex area. OBMC that uses the finite difference gradient calculation is highly dependent on the accuracy of the sensitivity information.

Table 6-5 shows the calibrated results derived using analytical sensitivity

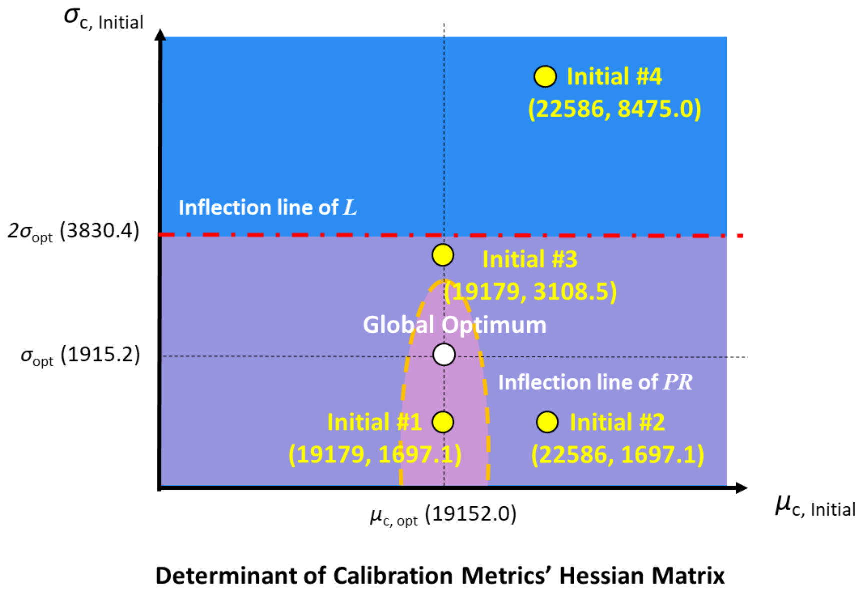

Figure 6-4 Four Initial Points and the Global Optimum on a Two-Dimensional Response Surface of Determinant of Calibration Metrics' Hessian Matrix

information. Compared with the previous case, the calibrated results show better performance. In particular, PR produces almost perfect results that are very close to the expected results. The calibrated results that come from using the likelihood function also provide much improved results; however, the OBMC starting from initial point #4 still gives calibrated results with low accuracy and robustness.

6.2.2.2. Subcase 2: Sequential Optimization and Uncertainty Propagation for Optimization-Based Model Calibration

The calibrated results from two different initial points are summarized in Table 6-6 and Table 6-7. Table 6-6 and Table 6-7 provide calibrated results of the unknown input variable (c_s) and predicted statistical moments (M_{Ypre}) of system responses ($q_{u, pre}$), with the objective function value (calibration metric) and the computational times in seconds. The two initial points are selected based on the analytical information derived in Section 3.3.1, such as the inflection separating the convex and non-convex area of the response surface of calibration metrics with respect to calibration parameters. Specifically, initial point #1 lies on the convex of the PR and initial point #2 lies on the non-convex of the PR. The two initial points and the global optimum are depicted in Figure 6-5. For the subcase in Section 6.2.2.2, the unknown input variable (Table 6-2), the soil cohesion (c_s), and the experimental observations (Table 6-3), bearing capacity (q_u), two cases are designed to check the effectiveness of the proposed method. The first case for initial point #1 presents a smaller value of standard deviation for the unknown input variable; the second case for initial point #2 presents a larger value of standard deviation for the unknown input variable.

From the initial point #1 (nonconvex area), 1) single-loop only with direct MCS and 2) single-loop only with UDR were used for calibrating the statistical parameters ($\mu_{c,\mathrm{opt}}$, $\sigma_{c,\mathrm{opt}}$) of the unknown input variable (c_s). The calibrated result derived from method 1 shows non-convergence of the optimization. This is because initial point #1 is on the nonconvex area when OBMC is conducted using PR. On the other hand, the calibrated result derived from method 2 shows convergence, because OBMC with the MMM formulates a convex optimization problem. However, the UP result shows numerical errors in the third (100.00%) statistical moment of the predicted system response. This numerical error led to an inaccurate calibrated result $\sigma_{c,\,\mathrm{opt}}$ (12.13%). The objective value of the calibration metric (MMM) also shows ‘1.00’, of which the value should be ‘0’ if it was accurately calibrated.

From initial point #2 (convex area), 1) single-loop only with direct MCS, 2) single-loop only with UDR, and 3) the proposed SOUP method were used to calibrate the statistical parameters ($\mu_{c,\,\mathrm{opt}}$, $\sigma_{c,\,\mathrm{opt}}$) of the unknown input variable (c_s). Since the optimization starts from inside of the convex area, the calibrated result from method 1 shows convergence and no significant numerical errors. To check the effect of numerical noise in the direct MCS, model calibration was carried out three times in succession. The repeated test results show that the numerical noise from the randomness did not affect the robustness of the OBMC. The calibrated result derived from method 2 also shows convergence. However, the UP result shows numerical errors in the third (100.00%) and fourth (29.66%) statistical moments of the predicted system responses. This numerical error led to an inaccurate calibrated result, $\sigma_{c,\,\mathrm{opt}}$ = 12.13% and a nonzero objective function, MMM = 1.10. The large value of standard deviation in the input variables ($\sigma_{c,}$ = 9576.2) resulted in larger numerical errors in the predicted statistical parameters of system responses, as compared to the results from case #1 (smaller std. ($\sigma_{c,}$ = 1915.2)) from the initial point #1. The calibrated result from method 3 shows a decrease in computational time compared to the calibrated result from method 1, and a decrease in numerical error compared to the calibrated result from method 2. The results from the case study prove that the proposed method offers improvements both in efficiency and accuracy.

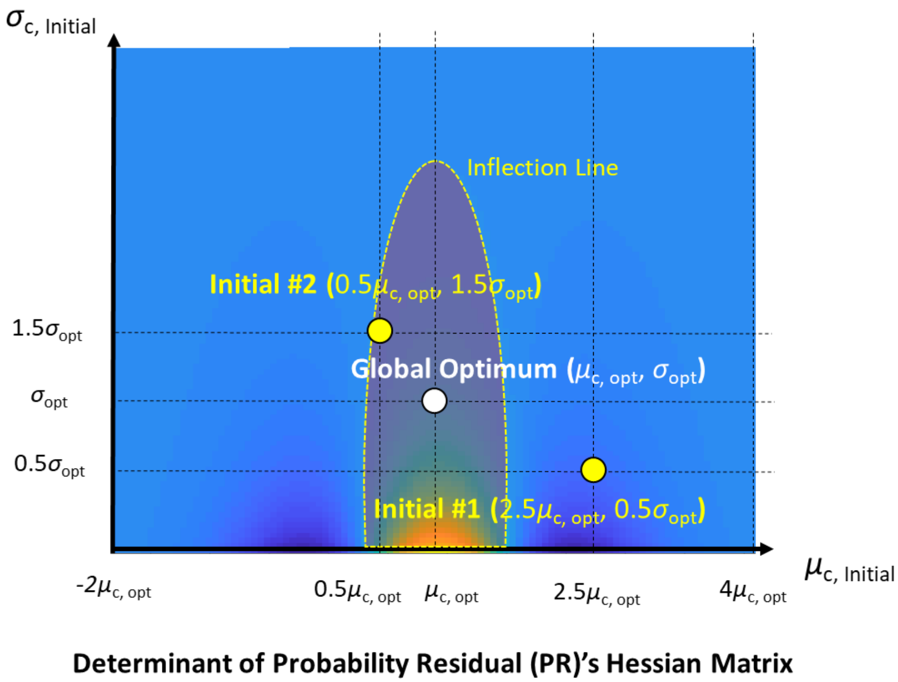

Figure 6-5 Two Initial Points and the Global Optimum on a Two-Dimensional Response Surface of Determinant of the Calibration Metrics' Hessian Matrix

Table 6-2 Properties of Input Variables in Terzaghi's Bearing Capacity Equation

Variable	**Unit**	**Type**	**Mean**	**Std. Dev.**
N_c	Dimensionless	Deterministic	17.7	-
N_q	Dimensionless	Deterministic	7.4	-
N_γ	Dimensionless	Deterministic	5.0	-
H_s	m	Deterministic	1.0	-
B_s	m	Deterministic	1.5	-
γ_s	kg/m^3	Normal	1842.0	184.2
c_s (Case #1)	kg/m^2	Normal	Unknown (19152.0)	Unknown (1915.2)
c_s (Case #2)	kg/m^2	Normal	Unknown (19152.0)	Unknown (9576.2)

Table 6-3 Statistical Information of Experimental Observations in Terzaghi's Bearing Capacity Example

System Response	**Unit**	**Type**	**Mean**	**Std. Dev.**
q_u (Case #1)	kg/m^2	Normal	359528.7	33933 (Small Std.)
q_u (Case #2)	kg/m^2	Normal	359528.7	254700 (Large Std.)

Table 6-4 Subcase 1: Calibrated Results from Optimization-Based Model Calibration without Analytical Sensitivity Information

	Likelihood Function		Probability Residual	
Initial Points	**Calibrated Mean**	**Calibrated Std.**	**Calibrated Mean**	**Calibrated Std.**
I (19179, 1697.1)	19145	1919.7	19143	1912.9
	19146	1915.5	19145	1913.3
	19165	**1860.4**	19030	1906.7
II (22586, 1697.1)	19141	1914.4	**19438**	**1779.0**
	18959	**3679.9**	**19629**	**1719.4**
	19133	**1747.5**	**17646**	**2077.1**
III (19179, 3108.5)	19142	1912.2	19146	1912.4
	18148	**2652.3**	**16823**	**2021.5**
	19129	**1894.4**	**26628**	**2630.6**
IV (22586, 8475.0)	**19123**	**3796.1**	19124	1909.5
	19151	**525.25**	19167	1908.5
	18439	**9659.6**	**19347**	**1890.8**

Table 6-5 Subcase 1: Calibrated Results from Optimization-Based Model Calibration with Analytical Sensitivity Information

	Likelihood Function		Probability Residual	
Initial Points	**Calibrated Mean**	**Calibrated Std.**	**Calibrated Mean**	**Calibrated Std.**
I (19179, 1697.1)	19155	1913.0	19153	1912.2
	19154	1913.3	19156	1912.4
	19154	1915.3	19152	1914.0
II (22586, 1697.1)	19153	1913.5	19150	1911.9
	19149	1914.6	19150	1913.2
	19153	1913.7	19152	1912.9
III (19179, 3108.5)	19154	1913.1	19154	1914.1
	19152	1912.7	19155	1913.7
	19152	1913.1	19154	1913.9
IV (22586, 8475.0)	**19530**	**70715**	19152	1912.5
	18967	**90523**	19150	1914.0
	19681	**70815**	19154	1913.8

Table 6-6 Subcase 2: Calibrated Results from Initial Point #1

	Calibrated Results		Obj.	Time (s)	Predicted Statistical Moments of System Response (Y_{pre})			
	Calibrated Mean	Calibrated Std.			M_1 ($\times 10^5$)	M_2 ($\times 10^4$)	M_3 ($\times 10^{-1}$)	M_4 ($\times 10^0$)
True Value	19152	1915.2	-	-	3.595	3.797	1.057	3.016
Single OUU by MCS	The OBMC diverges							
Avg. Error (%)	-	-	-	-	-	-	-	-
Single OUU by UDR	19153	1915.7	1.00	10.17	3.595	3.797	0.000	3.000
Avg. Error (%)	0.00	0.03	-	-	0.00	0.00	100.00	0.53

Table 6-7 Subcase 2: Calibrated Results from Initial Point #2

	Calibrated Results		Obj.	Time (s)	Predicted Statistical Moments of System Response (Y_{pre})			
	Calibrated Mean	**Calibrated Std.**			**M_1 ($\times 10^5$)**	**M_2 ($\times 10^4$)**	**M_3 ($\times 10^{-1}$)**	**M_4 ($\times 10^0$)**
True Value	19152	9576.2	-	-	3.595	2.547	8.897	4.265
Single OUU by MCS	19112	9587.7	0.00	95.64	3.589	2.540	8.866	4.246
	19161	9597.5	0.00	108.29	3.599	2.549	8.919	4.268
	19114	9592.6	0.00	103.04	3.588	2.543	8.879	4.240
Avg. Error (%)	0.45	0.49	-	-	0.47	0.51	0.80	1.10
Single OUU by UDR	19153	10738	1.10	10.17	3.595	2.543	0.000	3.000
Avg. Error (%)	0.00	12.13	-	-	0.00	0.16	100.00	29.66
SOUP by UDR & MCS	19154	9576.7	0.00	55.74	3.592	2.545	8.899	4.272
	19158	9574.3	0.00	57.25	3.594	2.541	8.844	4.269
	19158	9525.4	0.00	53.65	3.596	2.534	8.813	4.234
Avg. Error (%)	0.07	0.56	-		0.13	0.82	1.56	1.10

6.3 Case Study 3: Model Calibration of an Automobile Steering Wheel Column's Vibration Analysis Model

Section 6.3 introduces an example that studies a calibration problem of an automobile steering wheel column's vibration analysis model. The case study is formulated to check the applicability of the proposed ideas in Research Thrust 1, 3, and 4 to a real-world application. Section 6.3.1 presents the problem description of Case Study 3. Section 6.3.2 presents the uncertainty characterization of observed natural frequency considering measurement errors. Section 6.3.3 provides the calibrated results from the SOUP method proposed in Research Thrust 4, with the results in the previous Section 6.3.2. However, after the validity check by the proposed POC, the calibrated results by the first round of model calibration does not show the high credibility. Therefore, this study conducts model refinement and re-calibration process.

6.3.1 Problem Description

An automobile steering system is the collection of components that allows the driver to guide an automobile vehicle (Figure 6-6) (Pak et al. 1991) (Zaremba et al. 1998) (Demers 2001). The collection broadly include a steering wheel, a steering column, and a cowl cross bar. An automobile steering system transmits the vibration generated by the engine directly to the driver. Thus, a design for avoiding resonance of the vibration transmitted to the driver is desired. For achieving a resonance avoidance design, the natural frequency of an automobile steering system is an important system performance or response. This dissertation aims to calibrate an automobile's column model in a statistical sense.

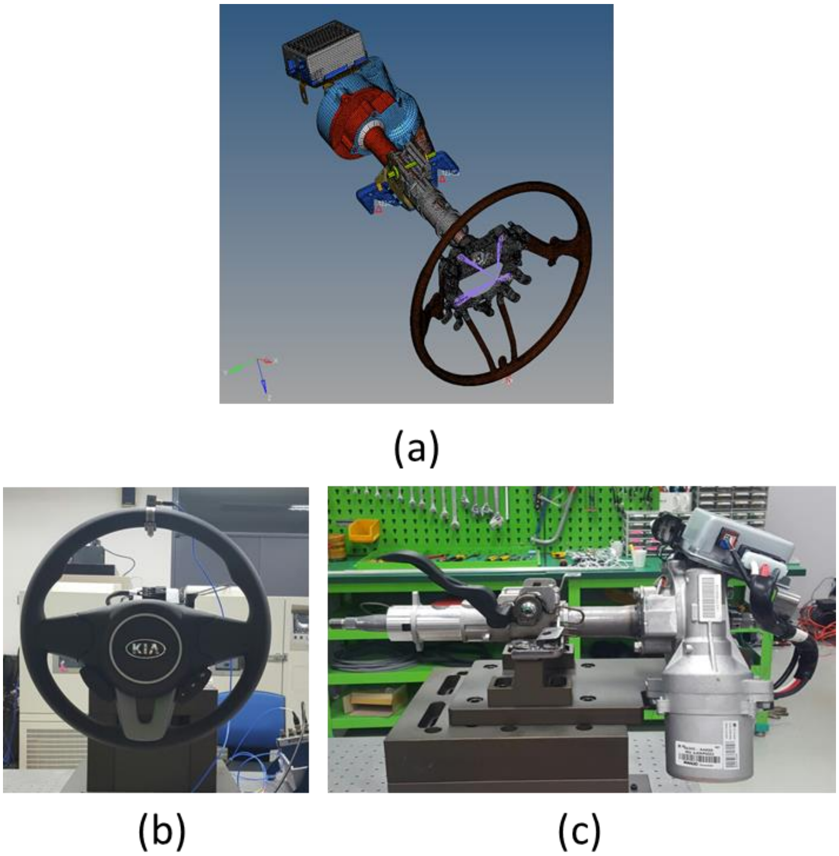

Figure 6-6 An Automobile Steering System including a Steering Wheel and a Steering Column: a) Finite Element Model of a Steering Wheel and a Steering Column, b) a Steering Wheel, and c) a Steering Column

6.3.2 Uncertainty Characterization of the Observed Natural Frequency of an Automobile Steering Column and Wheel Under Consideration of Measurement Errors

Three target vibrational modes are 1) vertical, 2) horizontal, and 3) rim modes, as depicted in Figure 6-7. Figure 6-7 presents the loading site by an impact hammer and the measuring site by a 3-axis acceleration sensor for three vibrational modes. The corresponding simulated vibration modes for each target vibration modes are the 1st, 2nd, and 4th natural frequency. (The 3rd natural frequency corresponds to a torsional vibration.)

An experiment for model calibration is designed for considering physical uncertainties and measurement errors as summarized in Table 6-8. First, three different steering wheels and six different steering columns are prepared to be used in experiments. Overall 18 combinations of steering wheel-columns can provide the

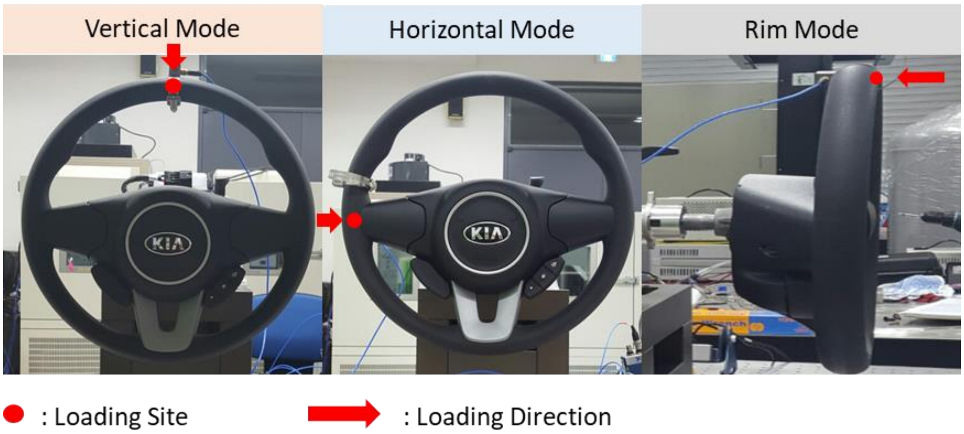

Figure 6-7 Three Vibrational Modes of a Steering Wheel Column: Vertical, Horizontal, and Rim Modes

variability by the physical uncertainties in manufacturing tolerance and material properties. Second, the experiments are conducted in three different experimental conditions (Laboratory at 1) Seoul national university, 2) Ajou university, and 3) Dong-Eui university). Three different experimental sites suggests systematic measurement errors. Each systematic measurement errors from three different experimental conditions could be quantified by the difference between the overall mean value and three mean values of each observation set from three experimental conditions. Third, measurements were repeated 5 times for each combination of steering wheel-column and experimental condition. With repetitions, random measurement errors could be quantified. Table 6-9 summarizes the quantified systematic and random measurement errors. Figure 6-8 presents the experimental observations of three vibrational modes. The proposed uncertainty characterization method that considers systematic and random measurement errors estimates the mean and standard deviation describing the true variability of three vibrational models (Table 6-10).

Table 6-8 Statistical Experiments for Model Calibration Considering Physical Uncertainties and Measurement Errors

Sources of Uncertainty and Error		**Numbers**	**Note**
Physical Uncertainties in Geometry and Material Properties	Variability in Steering Wheels	×3	Variability by Physical Uncertainties
	Variability in Columns	×6	Variability by Physical Uncertainties
Experimental Conditions	SNU/Ajou/Dong-Eui	×3	Systematic Measurement Error
Measurement Repetition		×5	Random Measurement Error
		Total 270	

Table 6-9 Quantification of Systematic and Random Measurement Errors

Measurement Error	Experimental Conditions	Degree of Error (%)	Note
Systematic Measurement Error	SNU	-0.19	Negative
	Ajou	-0.43	Negative
	Dong-Eui	0.74	Positive
Random Measurement Error	SNU	0.27/0.14/0.10	Vertical/Horizontal/Rim
	Ajou	0.32/0.37/0.24	
	Dong-Eui	0.11/0.11/0.07	

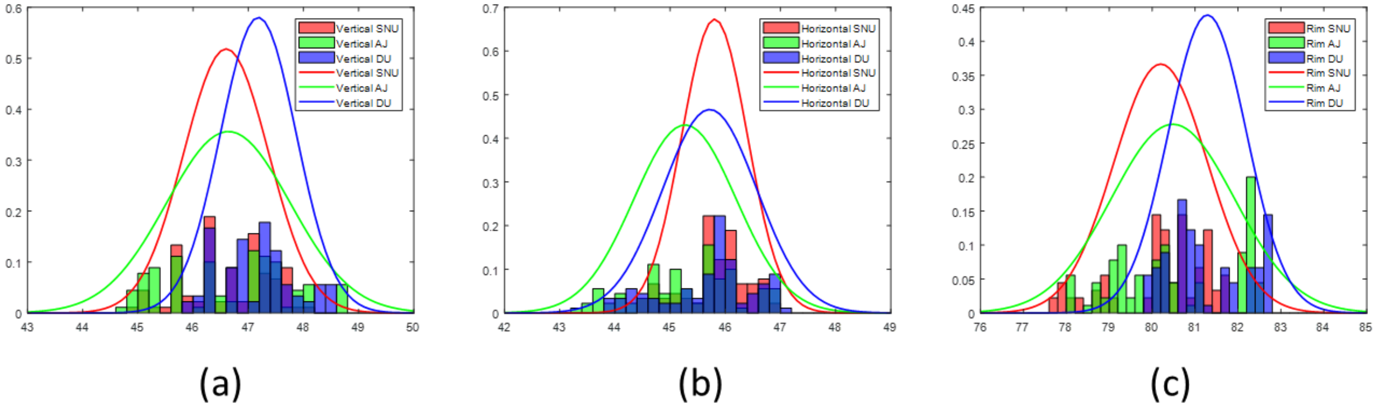

Figure 6-8 Experimental Observations of Three Vibrational Modes: a) Vertical, b) Horizontal, and 3) Rim Modes

Table 6-10 Characterized Uncertainty in Observed Three Vibrational Modes After Consideration of Measurement Errors

Experimental Observations	Unit	Number of Data	Mean	Standard Deviation	Note
Natural Freq. of Vertical Mode	Hz	90×3	46.8274	0.9136	QOI for Calibration
Natural Freq. of Horizontal Mode	Hz	90×3	45.6175	0.8338	QOI for Calibration
Natural Freq. of Rim Mode	Hz	90×3	80.6908	1.2479	QOI for Validation

6.3.3 1st Round of Model Calibration and Validation of Steering-Column Vibrational Model

This section presents the first round of model calibration and validation of steering-column vibrational model. First, the calibration is conducted by the proposed SOUP in Section 6.3.3.1. Second, the proposed POC provides the first round of validity check of the calibrated result.

6.3.3.1. Model Calibration by the Proposed SOUP

To improve the credibility of the automobile steering column model, a model calibration is formulated with two system responses (natural frequencies of vertical and horizontal modes) and two unknown input variables. Calibrating two unknowns by two known outputs formulates a determined problem. Among 36 input variables of the model (9 variables in the wheel part and 27 variables in the column part), the two unknown input variables are determined by a variable screening process: 1)

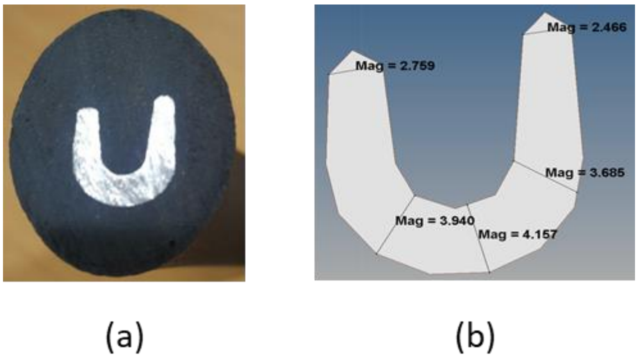

Figure 6-9 The Frame and Cover of a Steering Wheel: a) Cross Section of a Real Specimen, b) Shell Element in Finite Element Model

density of aluminum in the wheel part ($\rho_{\text{wheel, alum}}$) and 2) thickness of polyurethane in the wheel cover part ($t_{\text{wheel, poly}}$). It is acceptable choice of unknown input variables since there was an awkward modeling of the steering wheel frame (aluminum) and cover (polyurethane) as depicted in Figure 6-9. To substitute the expensive computational model, a Kriging surrogate modeling technique (Kleijnen 2009) is devised with Latin hyper cube design of experiments (Baghdasaryan et al. 2002) (Lee and Chen 2009). In this model calibration problem, the density of aluminum in the wheel part ($\rho_{\text{wheel, alum}}$) and the thickness of polyurethane in the wheel cover part ($t_{\text{wheel, poly}}$) are assumed to follow a normal distribution. Using the information for the system response observed in Section 6.3.2, model calibration is conducted using the SOUP method proposed in Research Thrust 3. (However, it is expected that the low degree of systematic and random error quantified in Section 6.3.2 may not impact the significant change in the calibrated result.) Table 6-11 presents the calibrated results by OBMC with considering measurement errors and using the proposed SOUP method.

6.3.3.2. Model Validation by the Proposed POC

To validate the calibrated result in previous section, a new validation metric, POC, is adopted. Using the area metric with hypothesis testing for model validation, however, is impossible, because the characterized variability in Section 6.3.2 is a probability distribution. With a parametric probability distribution, an area metric, which is based on the empirical CDF of experimental observation is not able to be calculated. In the end, in this case study, the proposed POC is used for validating the calibration result.

The model validation is conducted with one system response natural frequency which corresponds to the rim vibration mode. The statistical information of characterized uncertainty in observations of the rim vibration mode is summarized in Table 6-10. The predictions are characterized by a nonparametric distribution. Thus, two probability distributions are compared by POC, calculated as ’76.50 %.’ The POC value before calibration is calculated as ‘16.51%.’ This results shows that the model calibration improved the credibility of the model credibility by ’76.50 %’. However, the POC value shows the low predictive credibility of the calibrated model. Thus, it can be concluded that the calibrated model requires model improvement or refinement process (Oh et al. 2016).

6.3.4 2nd Round of Model Calibration and Validation of Steering-Column Vibrational Model

This section presents the second round of model calibration and validation of steering-column vibrational model. First, the calibration problem is reformulated and the unknown input variables are calibrated by the proposed SOUP. Second, the proposed POC provides the second round of validity check of the calibrated result.

After a systematic model refinement following the previous study (Oh et al. 2016), it is found that there exist awkward modeling of the linkage element between a steering wheel and air bag (Figure 6-10). This study modify the rigid link to a spring element, of which stiffness is set to new calibration parameter. Therefore, in the 2nd round of model calibration and validation, the two unknown input variables are determined: 1) thickness of polyurethane in the wheel cover part ($t_{\text{wheel, poly}}$) and 2) stiffness of air bag-wheel spring linkage. Table 6-12 presents the calibrated results

(a) (b)

Figure 6-10 (a) Steering Wheel and Air Bag; (b) Finite Element Model of the Steering Wheel

by OBMC with considering measurement errors and using the proposed SOUP method. The proposed POC is used for validating the calibration result. The validation result shows the improvement in predictive capability that is increased from 76.5% to 96.5%. In the end, using the calibrated model, it is expected that more accurate prediction will be available.

Table 6-11 The First Round Calibrated Results of Unknown Input Variables of the Steering Wheel Column Vibrational Model

Variable	**Unit**	**Type**	**Before Calibration**			**Calibrated Results**		
			Mean	**Standard Deviation**	**COV**	**Mean**	**Standard Deviation**	**COV**
$\rho_{\text{wheel, alum}}$	g/cm^3	Normal	2.76	0.14	5.07%	2.50	0.14	5.60%
$t_{\text{wheel, poly}}$	mm	Normal	5.00	0.25	5.00%	9.76	1.08	11.13%

Table 6-12 The Second Round Calibrated Results of Unknown Input Variables of the Steering Wheel Column Vibrational Model

Variable	**Unit**	**Type**	**Before Calibration**			**Calibrated Results**		
			Mean	**Standard Deviation**	**COV**	**Mean**	**Standard Deviation**	**COV**
k_{air}	N/mm	Normal	1000	100.0	10.00%	1071	332.5	31.05%
$t_{\text{wheel, poly}}$	mm	Normal	9.76	1.08	11.13%	3.02	1.29	42.72%

6.4 Case Study 4: Model Calibration of a Thin-Film Transistor Liquid Crystal Display Panel Deflection Model

Section 6.4 introduces an example that studies a calibration problem of a thin-film transistor liquid crystal display (TFT-LCD) panel's deflection model. The case study is formulated to check the applicability of the proposed ideas in Research Thrust 1 and 3 to a real-world application. Section 6.4.1 presents the problem description of Case Study 4. Section 6.4.2 presents the uncertainty characterization of observed deflection from two different experimental conditions. Section 6.4.3 provides the calibrated results from the SOUP method proposed in Research Thrust 3, with the results in the previous Section 6.4.2. At the end, Section 6.4.4 provides the validity check of calibrated results from Section 6.4.3.

6.4.1 Problem Description

A TFT-LCD is a popular flat-panel display that is used in appliances such as televisions, monitors, mobile phones, and others. A TFT-LCD adopts Thin-film transistor (TFT) technology as its active switching device (Katayama 1999) (Kim et al. 2004) (Pan and Chen 2007) (Vepakomma et al. 2013). Figure 6-11 describes the structure of a TFT-LCD, which consists of polarizers, a color filter, glass substrates, a thin-film transistor array, a back light unit, and a thin layer of liquid crystals (Su et al. 2012). To manufacture a certain size of TFT-LCD panel for various types of appliances, a one drop fill liquid crystal injection process is used, and a large mother panel is cut into medium/small panel sizes by 1) diamond wheel cutting or 2) laser

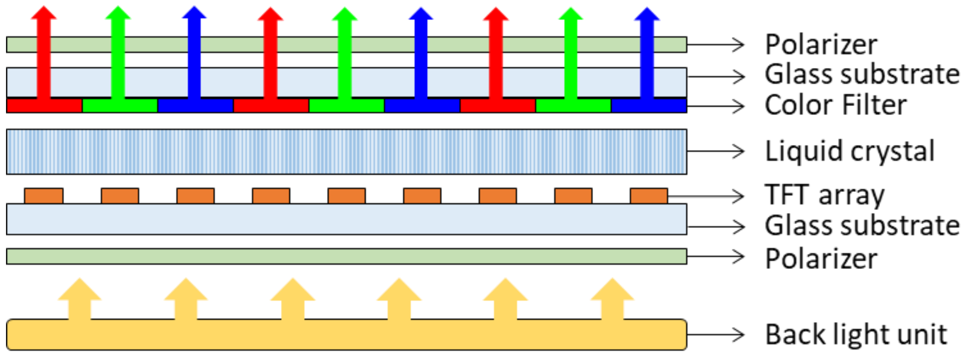

Figure 6-11 The Scheme of a Thin-Film Transistor Liquid Crystal Display

cutting (Figure 6-12) (Pan et al. 2008). The method of diamond wheel cutting is a mechanical way of cutting the mother glass panel, which includes scribing the glass using a wheel and cleaving the substrate using a mechanical stress. Due to the thick scribing on the panel by contact cutting and micro cracks that can develop during the cleaving, large variations (manufacturing tolerance or error) exist in panels that are manufactured by the diamond wheel cutting process. On the other hand, laser cutting utilizes an ultraviolet laser for the scribing the mother glass by noncontact cutting. This leads to small variations in manufacturing quality. The difficulties in adopting laser cutting include the requirement of a pre-bending technique and the need to control the thermal stress (Tsai and Lin 2007). To apply the pre-bending technique, a computational model is used to predict deflection along the cutting path.

Electric appliances, such as mobile phones, are subjected to various mechanical loads during transportation, dropping, or bending while in use (Chung et al. 2011). In general, to examine mechanical failures, 3-point or 4-point bending experiments and simulations are applied (Cui and Wisnom 1992) (Quinn et al. 2009) (Lin et al. 2010) (Hein and Brancheriau 2018). (Drop simulation and experiments are available, however, they have difficulties in repetition of experimental observations and validation due to the many uncontrollable factors inherent in these types of experiments and simulations. (Pan and Chen 2007))

This case study aims to calibrate a computational model for a TFT-LCD panel with 3-point bending experiments. The calibrated and validated model with 3-point bending performance can be used for 1) examination of future failure by transportation, or dropping and bending while in use, and 2) design of laser cutting manufacturing by a pre-bending technique. For accurate model calibration, uncertainty characterization, as described in Research Thrust 1, and OBMC by SOUP, as proposed in Research Thrust 3, are implemented in Section 6.4.2 and Section 6.4.3, respectively.

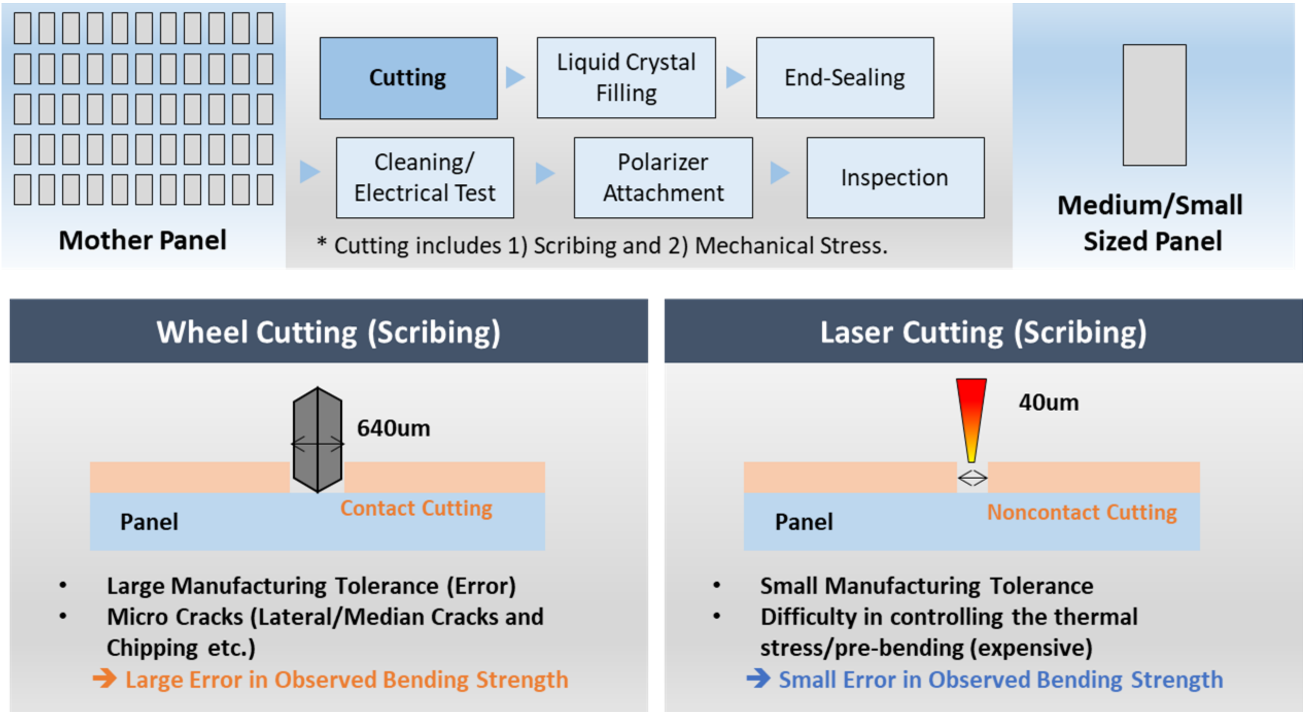

Figure 6-12 The Manufacturing Process of Cutting an LCD Panel: Wheel and Laser Cutting

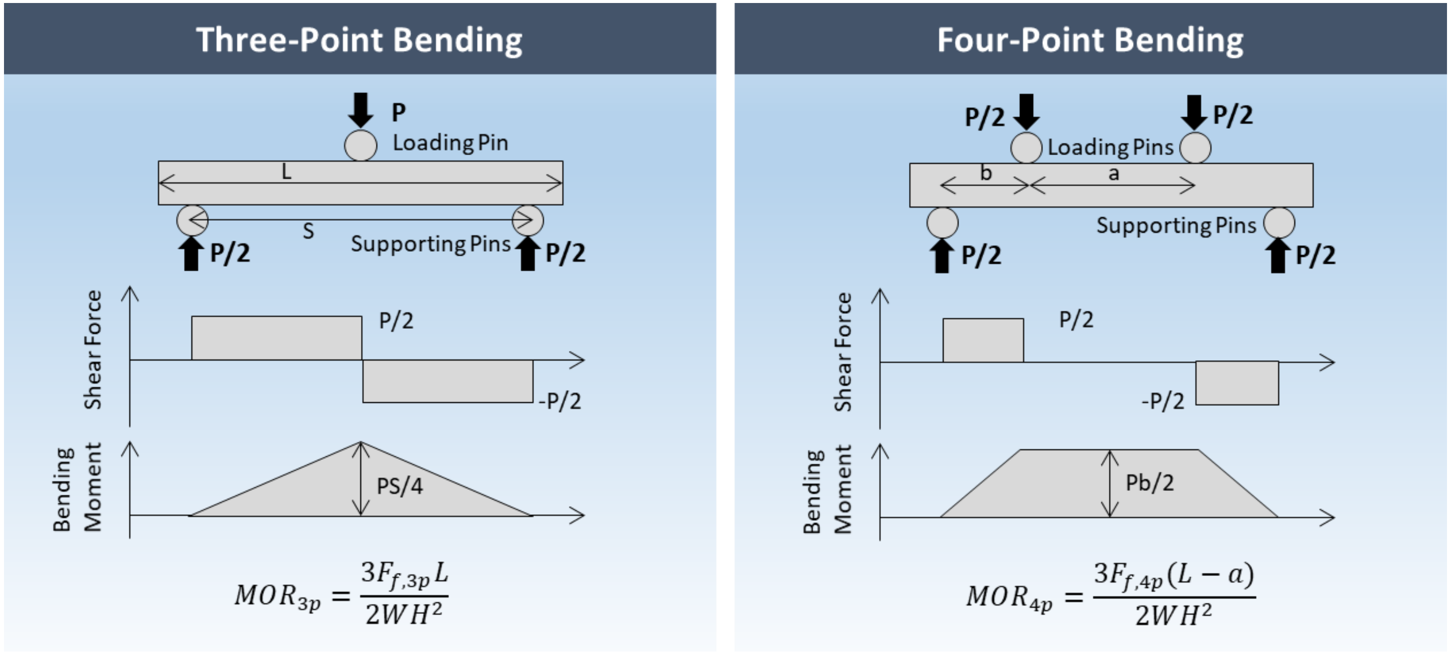

Figure 6-13 Three-Point/Four-Point Bending Tests of Brittle Materials

6.4.2 Uncertainty Characterization of Observed Deflection of a Thin-Film Transistor Liquid Crystal Display Panel's Deflection from Different Experimental Conditions

A 3-point or 4-point bending test is a general way to evaluate the quality of a glass panel by measuring the static bending strength (Figure 6-13). With a 3-point or 4-point bending test, a force-displacement diagram can be obtained when a force is applied in three directions: Top X, Top Y, and Bottom X (Figure 6-14). It is known that a 3-point bending test provides higher bending strength than the bending strength obtained by a 4-point bending test. In Figure 6-13, it can be confirmed that a larger volume of material is subjected to the same bending moment in a four-point bending. A failure from bending stress begins from a micro crack produced from the manufacturing process, which means the larger the volume is, the higher the probability that a defect exists.

Table 6-13 presents experimental observations of the bending strength from two different experimental and manufacturing conditions: 1) 3-point bending/diamond wheel cutting, and 2) 4-point bending/laser cutting. Three different system responses are observed under two different experimental and manufacturing conditions, as depicted in Figure 6-14: 1) Top-X (the front face of the panel facing upward and jig

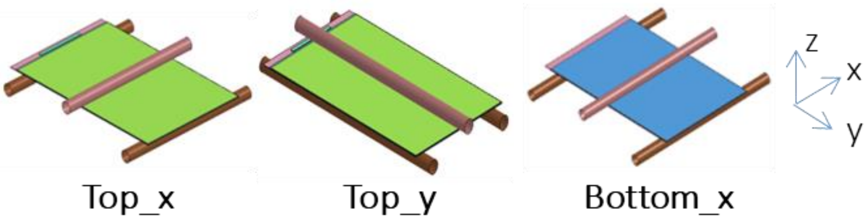

Figure 6-14 Three-Point Bending Test of a TFT-LCD Panel in Three Directions

parallel to the x axis), 2) Top-Y (the jig parallel to the y axis), and 3) Bottom-X (the back face of the panel facing upward).

This case study aims to calibrate a computational model for two reasons: 1) examination of future failure that may result from transportation, or dropping or bending while in use, 2) design of the laser cutting manufacturing process through a pre-bending technique. Therefore, this case study calibrates a computational model about 3-point bending for the experimental condition and laser cutting for the manufacturing method. Thus, uncertainty characterization for experimental observations from different experimental conditions and manufacturing methods is required.

From the reference (Cui and Wisnom 1992), the use of a 4-point bending test for a glass substrate, rather than a 3-point bending test, reduces the bending strength by approximately 20%. The difference in the bending strength observed from the two different experimental conditions can be recognized as a systematic error. In this case, it can be considered that the 4-point bending test has 20 % negative systematic error. Also, the panel specimen manufactured by diamond wheel cutting has more random errors, as explained in Section 6.4.1, and the degree of random measurement errors is identified as 10% more, compared with a panel produced by the laser cutting method (Pan et al. 2008). Table 6-14 presents the results of uncertainty characterization using the method proposed in Research Thrust 1. As depicted in Figure 6-15, the probability distribution of QOI (Top-X Bending Strength) can be obtained.

Table 6-13 The Statistical Information of Experimental Observations from Different Experimental Conditions and Manufacturing Methods (Before Considering the Measurement Errors)

Number	**Bending Direction**	**# of Data**	**Bending Method**	**Cutting Method**	**Mean**	**Std.**	**Note**
#1	Top X	10	3-Point	Wheel	53.20	5.74	QOI for Calibration
#2	Top X	10	4-point	Laser	44.33	2.04	
#3	Top Y	10	3-point	Wheel	110.70	31.57	QOI for Calibration
#4	Top Y	10	4-point	Laser	92.25	4.25	
#5	Bottom X	10	3-point	Laser	31.57	2.27	QOI for Validation

Table 6-14 The Statistical Information of Experimental Observations from Different Experimental Conditions and Manufacturing Methods (After Considering the Measurement Errors)

Number	**Bending Direction**	**# of Data**	**Bending Method**	**Cutting Method**	**Mean**	**Std.**	**Note**
#1	Top X	20	3-Point	Laser	53.16	1.17	QOI for Calibration
#2	Top Y	20	3-Point	Laser	110.17	2.31	QOI for Calibration
#3	Bottom X	10	3-Point	Laser	31.57	2.27	QOI for Validation

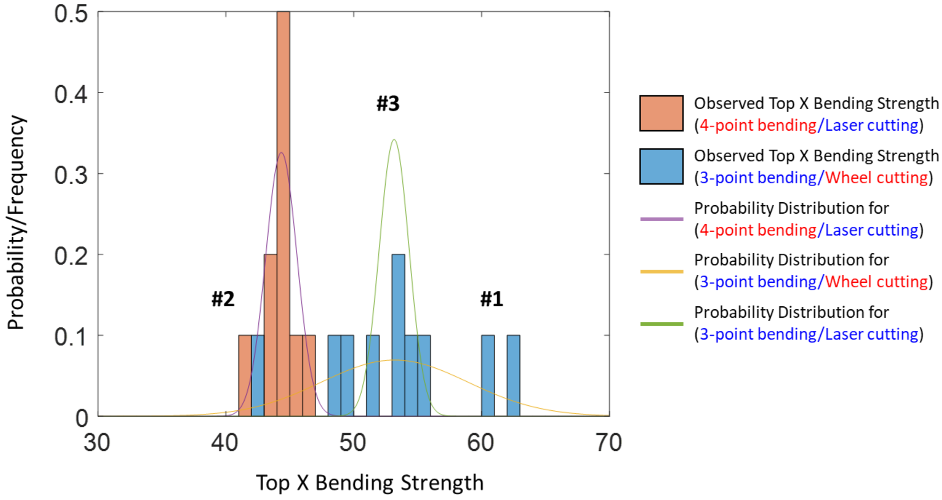

Figure 6-15 The Probability Distribution of the Experimental Observations

6.4.3 Model Calibration Using Sequential Optimization and Uncertainty Propagation

To improve the credibility of the TFT-LCD panel deflection model, a model calibration is formulated with two system responses (deflections by Top-X, Top-Y loadings) and two unknown input variables (the Young's modulus of glass (E_g) and polarizer (E_{py})). In this model calibration problem, the Young's modulus of glass (E_g) and polarizer (E_{py}) is assumed to follow a log-normal distribution and geometric properties are set to deterministic. Using the information for the system response observed in Section 6.4.2, model calibration is conducted using the SOUP method proposed in Research Thrust 3. To simulate 3-point bending, finite element analysis is performed using LS-DYNA. Due to the computational burden required by finite element analysis, Kriging surrogate modeling and perturbation by $\pm 3\sigma$ of each variable for design of experiments are used.

Table 6-10 presents the calibrated results. First, compared with the initial value of the mean and standard deviation of the unknown input variables, the calibrated mean and standard deviation are both decreased. Thus, using the calibrated value, more accurate prediction will be available. The column listing the results after calibration in Table 6-10 presents three cases of calibrated results: 1) OBMC without considering measurement errors, 2) OBMC with considering measurement errors, but not by SOUP, and 3) OBMC with considering measurement errors and using the proposed SOUP method. The first case is the result from the previous study (Jung 2011) (Oh et al. 2016). Since the previous study did not consider the existence of measurement errors, the calibrated results may be wrong. Unfortunately, the second case did not show the convergence of optimization by OBMC, due to the possible causes discussed in Research Thrust 2.

Table 6-15 The Calibrated Results of Unknown Input Variables of the Thin-Film Transistor Liquid Display Panel Deflection Model

Variable	Unit	Type	Before Calibration		After Calibration			Note
			Mean	Std. Dev.	Mean	Std. Dev.	C.V.	
E_g	GPa	Lognormal	93.77	4.29	72.13	2.86	3.97%	No Consideration of Errors
					No Convergence			Original OBMC
					72.20	2.30	3.19%	OBMC by SOUP
E_{py}	GPa	Lognormal	4.42	1.94	3.40	1.29	37.94%	No Consideration of Errors
					No Convergence			Original OBMC
					3.43	0.58	16.91%	OBMC by SOUP

6.4.4 Model Validation of Calibrated Results

To validate the calibrated results in Section 6.4.3, model validation is conducted with one system response (deflection by Bottom-X). Because the experimental observations of deflection by Bottom-X loading are obtained from an experimental condition where the 3-point bending test is used to measure the deflection and laser cutting method is used to manufacture the small size of panel. In other words, there is no need to use uncertainty characterization method that considers measurement errors. Thereby, the area metric with hypothesis testing is able to be used for validating the calibrated results. Case Study 4 compares the usage of the area metric with hypothesis testing (Section 6.3.4.1) with the proposed probability of coincidence (POC) (Section 6.3.4.2).

6.4.4.1. Model Validation by the Area Metric and Hypothesis Testing

For the validity check by the area metric and hypothesis testing, first, the area metric between u-value CDFs from simulated predictions ($F(u_{\mathrm{pre}})$) and 10 numbers of observations ($S(u_{\mathrm{obs}})$) is calculated as '0.0933', as shown in Figure 6-16. The critical value of the area metric at a significance level of 0.05 is provided as '0.1805' in case ten experimental data are used for the validity check as shown in Figure 6-17. The smaller value of the area metric '0.0933' than the critical value denotes that the null hypothesis which states a computational model is valid, would not be rejected. In conclusion, there is not enough evidence to conclude that the calibrated model is invalid. However, strictly speaking, this conclusion does not prove that the calibrated model is valid to be used for intended use.

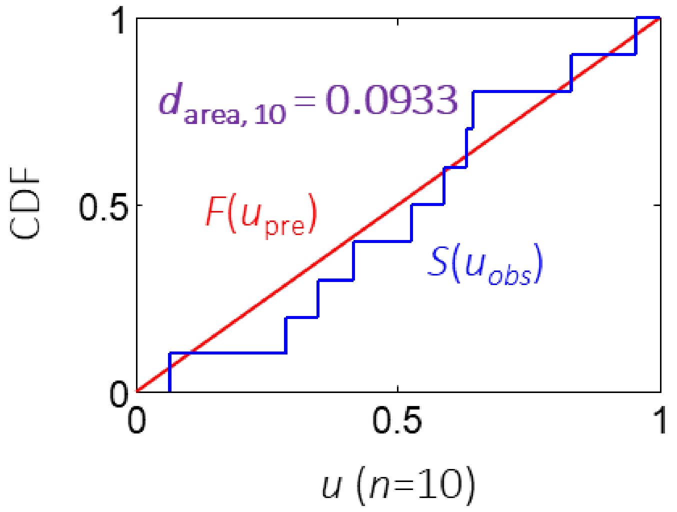

Figure 6-16 Calculation of the Area Metric by u-Value CDFs

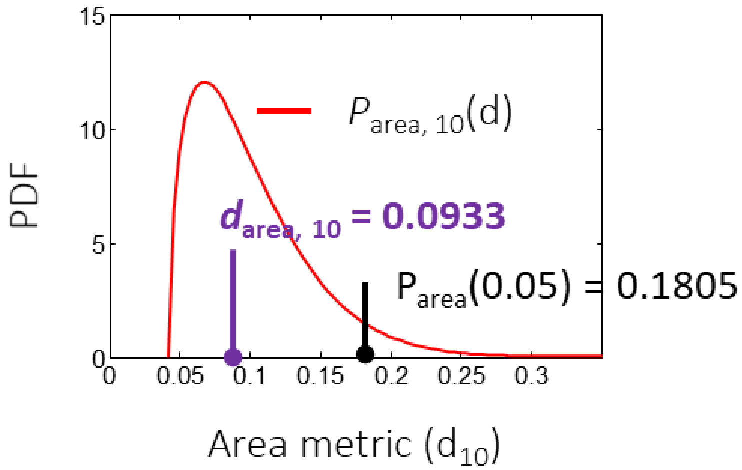

Figure 6-17 Probability Density Function of the Area Metric When the Number of Observations is Ten

6.4.4.2. Model Validation by the Probability of Coincidence

A new statistical validation metric, POC, is employed to quantify the probability degree of coincidence. To use POC, the observed deflections by a loading Bottom-X are assumed to follow a normal distribution. The predictions are characterized by a nonparametric distribution. Thus, two probability distributions are compared by POC, calculated as '98.03 %'. This results provides an analyst who will use the calibrated model that the model has the credibility by '98.03 %'.

Chapter 7

Conclusion

7.1 Summary and Contributions

The ultimate goal of this doctoral dissertation is to provide a process to improve the credibility of computational models by enabling calibration of the unknown input variables in computational models. The proposed research in this doctoral dissertation aims to establish a complete optimization-based model calibration (OBMC) method, by tackling four technical issues. It is expected that the proposed research offers the following potential contributions and broader impacts to work towards achieving a long-held vision of engineers to build highly credible computational models.

Contribution 1: Uncertainty Characterization of Experimental Observations using Parametric Probability Distributions Under Consideration of Measurement Errors

Research Thrust 1 in this doctoral dissertation proposes an uncertainty characterization method that uses maximum likelihood estimation. To best of the author's knowledge, most existing works in uncertainty characterization are focused on how to quantify the variability in a quantity of interest through a probability distribution while ignoring measurement errors. By applying the proposed method to case studies, this research confirmed that the proposed method can contribute to: 1) characterizing the accurate probability distribution by eliminating the effect of measurement errors and 2) easing the degree of statistical uncertainty present from a lack of observation data by integrating experimental observation data from different experimental conditions. Through these two contributions, an improved calibration by OBMC is available.

Contribution 2: Identification of Possible Causes of Inaccurate and Unstable Calibrated Results that occur for OBMC that uses the Existing Two Calibration Metrics and Improvement of OBMC with Analytical Sensitivity Information

Research Thrust 2 identifies the possible causes of inaccurate and unstable model calibration resulting from OBMC that uses either the likelihood function or probability residual. The analytical and numerical investigations reveal 1) non-convexity of OBMC when it is performed with either of the two existing calibration metrics, 2) a low degree of sensitivity and approximate calculation of sensitivity information, which is required to conduct gradient-based optimization, and 3) randomness due to sampling methods used for the uncertainty propagation process cause the inaccurate and unstable calibrated results. To address these issues, this dissertation research proposes the idea of OBMC with analytical sensitivity information to improve the calibration. Additionally, a comparative study of existing calibration metrics is provided for future use of these methods in model calibration. Investigation of the technical issues that should be addressed to achieve improved OBMC motivated the research outlined in Research Thrust 3.

Contribution 3: Development of Sequential Optimization and Uncertainty Propagation for Efficient and Accurate Model Calibration using OBMC

OBMC adopts an optimization under uncertainty (OUU) process that requires expensive computational cost. Research Thrust 3 proposes a process of sequential optimization and uncertainty propagation (SOUP) to improve the efficiency of OBMC, while retaining the accuracy. To conduct SOUP, a new calibration metric, the moment matching metric, is proposed, which makes an effort to address the technical issues introduced in Research Thrust 3. In addition, the review study that examines the developed OUU formulations and uncertainty propagation methods provides a guide to develop an OUU loop for OBMC.

Contribution 4: Proposition of a New Validation Metric - Probability of Coincidence

Research Thrust 4 presents a new validation metric, called probability of coincidence (POC), to statistically validate the calibrated results. To the best of the author’s knowledge, the area metric with hypothesis testing has shown best performance for a validation method, among available options. However, to the best of author's knowledge, the area metric with hypothesis testing still has limitations. Research Thrust 4 attempts to overcome these limitations, while satisfying the desired characteristics of a validation method. The POC is not a perfect validation metric, however, it is worth suggesting various ways of validating a computational model. In the end, the discussion provides the desired characteristics of a validation method for future work.

7.2 Suggestions for Future Research

Although the technical advances proposed in this doctoral dissertation successfully address four technical challenges in OBMC, there are still several research topics that should be explored, further investigations undertaken, and developments required to achieve a complete OBMC. Specific suggestions for future research are listed as follows.

Suggestion 1: A Systematic Way to Identify the Sources and Magnitude of Measurement Errors

Development of a systematic way to identify the sources and magnitude of measurement errors is a significant research need. While it is confirmed that the proposed uncertainty characterization that considers measurement errors is a promising method, its application can only be realized after the sources and magnitude of measurement errors are recognized. Future work should focus on recognition of the sources and magnitude of measurement errors that arise from designing to conducting the experiments. Additionally, a systematic way to integrate the experimental observations from different experiment conditions is needed.

Suggestion 2: Advanced Uncertainty Characterization Using Bayesian Approaches and Bound Information

From the result of Case Study 1, the estimated statistical parameters of the probability distribution of observed system responses show large variations even with large numbers (e.g., 50 observations in Figure 6-2) of experimental observations. This dissertation presents two possible future works using Bayesian approaches and bound information. First, the Bayesian approach estimates the statistical parameters' statistical parameters of the probability distribution of the quantity in interest. Thereby, the variations in the statistical parameters of the probability distribution of the system responses can be quantified using Bayesian methods, which can be helpful for understanding the variations in the uncertainties (Park et al. 2010) (Zhang et al. 2011). Second, providing bound information, such as upper and lower limits of the quantity in interest, can decrease the large variation of estimations (Kang et al. 2016) (Kang et al. 2018). In addition, though these two future works, the optimum number or minimum number of experimental observations can be provided for accurate characterization of uncertainties in the quantity of interest.

Suggestion 3: Sequential/Trust-Region Based Surrogate Modeling Methods

For now, the second sequence of the proposed SOUP method uses direct Monte Carlo simulation (MCS) for the uncertainty propagation (UP) process. In reality, implementing direct MCS, requires a significant number of computations, such as one million, to an already computationally expensive computational model is impractical. Even for the first sequence of SOUP, $(M - 1) \times N + 1$ (N is the number of input variables and M is the number of integration points along each random variable) number of computations for the univariate integration can be a computational burden. In this case, a surrogate modeling technique must be employed to substitute for the expensive computational model (Hill and Hunter 1966) (Montgomery 2008) (Kang et al. 2010) (Roussouly et al. 2013) (Myers et al. 2016). Thus, building an accurate surrogate model becomes an important issue. (Due to the effectiveness of the first sequence of the proposed SOUP, the optimization domain for the second sequence can be diminished.) Future work will thus include attempts to provide a breakthrough to enhance the accuracy of surrogate modeling (e.g., a sequential method (Du and Chen 2002) (van Beers and Kleijnen 2008) (Li et al. 2010) (Gorissen et al. 2010) (Yuan and Ng 2013) (Atamturktur et al. 2013), trust-region method (Giunta and Eldred 2000) (Zhou et al. 2007), deep learning (Goodfellow et al. 2016), or others).

Suggestion 4: Development of a New Validation Method to Address Limitations of the Area Metric and Proposed Probability of Coincidence

Even though the area metric with hypothesis testing is the most-used method in the current V&V field, there has also been attempts to develop other validation methods, for example, validation metrics based on the comparison of probability density functions (Rebba and Mahadevan 2008) (Xiong et al. 2009) (Harmel et al. 2010) (Kokkolaras et al. 2013) (Mullins et al. 2016), degree of overlap (Harmel et al. 2010), and validation methods based on the interval (Chen et al. 2004) (Buranathiti et al. 2006) (Ghanem et al. 2008) (Halder and Bhattacharya 2011). However, these other methods are limited in their ability to be used in practical applications. For example, a large number of experimental observation are required for their use. In the end, it can be said that a validation method that satisfies all desired characteristics presented in this dissertation has not yet been developed and, thus, future research efforts are needed.

References

Agarwal H, Renaud JE, Preston EL, Padmanabhan D (2004) Uncertainty quantification using evidence theory in multidisciplinary design optimization Reliability Engineering & System Safety 85:281-294 doi:10.1016/j.ress.2004.03.017

Anderson TW, Darling DA (1952) Asymptotic theory of certain" goodness of fit" criteria based on stochastic processes The annals of mathematical statistics:193-212

Anderson TW, Darling DA (1954) A test of goodness of fit Journal of the American statistical association 49:765-769

Aoues Y, Chateauneuf A (2010) Benchmark study of numerical methods for reliability-based design optimization Structural and multidisciplinary optimization 41:277-294

Arendt PD, Apley DW, Chen W (2012a) Quantification of model uncertainty: Calibration, model discrepancy, and identifiability Journal of Mechanical Design 134:100908

Arendt PD, Apley DW, Chen W, Lamb D, Gorsich D (2012b) Improving identifiability in model calibration using multiple responses Journal of Mechanical Design 134:100909

Arora JS (2004) Introduction to optimum design. Elsevier,

Atamturktur S, Williams B, Egeberg M, Unal C (2013) Batch sequential design of optimal experiments for improved predictive maturity in physics-based modeling Structural and Multidisciplinary Optimization 48:549-569

Babuska I, Oden JT (2004) Verification and validation in computational engineering and science: basic concepts Computer Methods in Applied Mechanics and Engineering 193:4057-4066

Baghdasaryan L, Chen W, Buranathiti T, Cao J Model validation via uncertainty propagation using response surface models. In: ASME 2002 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, 2002. American Society of Mechanical Engineers, pp 981-992

Bayarri MJ et al. (2007) A framework for validation of computer models Technometrics 49

Benek JA, Kraft EM, Lauer RF (1998) Validation issues for engine-airframe integration AIAA journal 36:759-764

Beyer H-G, Sendhoff B (2007) Robust optimization – A comprehensive survey Computer Methods in Applied Mechanics and Engineering 196:3190-3218 doi:10.1016/j.cma.2007.03.003

Borg A, Paulsen Husted B, Njå O (2014) The concept of validation of numerical models for consequence analysis Reliability Engineering & System Safety 125:36-45 doi:10.1016/j.ress.2013.09.009

Boyd S, Vandenberghe L (2004) Convex optimization. Cambridge university press,

Buranathiti T, Cao J, Chen W, Baghdasaryan L, Xia ZC (2006) Approaches for model validation: methodology and illustration on a sheet metal flanging process Journal of Manufacturing Science and Engineering 128:588-597

Campbell K (2006) Statistical calibration of computer simulations Reliability Engineering & System Safety 91:1358-1363

Charnes A, Frome E, Yu P-L (1976) The equivalence of generalized least squares and maximum likelihood estimates in the exponential family Journal of the American Statistical Association 71:169-171

Chen W, Baghdasaryan L, Buranathiti T, Cao J (2004) Model validation via uncertainty propagation and data transformations AIAA journal 42:1406-1415

Choi H, Jung JH, Oh H, Youn BD Statistical Validation for Rotor Dynamics Model of a Journal Bearing Rotor System. In: presented in The 12th World Congress on Computational Mechanics, Seoul, Republic of korea, July 24-29 2016.

Choi J, An D, Won J (2010) Bayesian approach for structural reliability analysis and optimization using the Kriging Dimension Reduction Method Journal of Mechanical Design 132:051003

Chung S-W, Jeong J-W, Oh S, Moon Y-J, Baek I-Y, Hong D-S Drop reliability of glass panel for LCD. In: Electronic Components and Technology Conference (ECTC), 2011 IEEE 61st, 2011. IEEE, pp 1250-1254

Cui WC, Wisnom MR (1992) Contact finite element analysis of three-and four-point short-beam bending of unidirectional composites Composites science and technology 45:323-334

da Silva Hack P, Schwengber ten Caten C (2012) Measurement Uncertainty: Literature Review and Research Trends IEEE Transactions on Instrumentation and Measurement 61:2116-2124 doi:10.1109/tim.2012.2193694

De Cursi ES, Sampaio R (2015) Uncertainty quantification and stochastic modeling with matlab. Elsevier,

Demers MA (2001) Steering wheel vibration diagnosis. SAE Technical Paper, Du X, Chen W Sequential optimization and reliability assessment method for efficient probabilistic design. In: ASME 2002 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, 2002. American Society of Mechanical Engineers, pp 871-880

Easterling RG (2001) Measuring the predictive capability of computational methods: Principles and methods, issues and illustrations SAND2001-0243, Sandia National Laboratories

Eldred MS, Giunta AA, Wojtkiewicz S, Trucano TG Formulations for surrogate-based optimization under uncertainty. In: 9th AIAA/ISSMO symposium on multidisciplinary analysis and optimization, Atlanta, GA, 2002.

Enevoldsen I, Sørensen JD (1994) Reliability-based optimization in structural engineering Structural safety 15:169-196

Fender J, Duddeck F, Zimmermann M (2014) On the calibration of simplified vehicle crash models Structural and Multidisciplinary Optimization 49:455-469

Ferson S, Ginzburg LR (1996) Different methods are needed to propagate ignorance and variability Reliability Engineering & System Safety 54:133-144

Ferson S, Joslyn CA, Helton JC, Oberkampf WL, Sentz K (2004) Summary from the epistemic uncertainty workshop: consensus amid diversity Reliability Engineering & System Safety 85:355-369

Ferson S, Oberkampf WL (2009) Validation of imprecise probability models International Journal of Reliability and Safety 3:3-22

Ferson S, Oberkampf WL, Ginzburg L (2008) Model validation and predictive capability for the thermal challenge problem Computer Methods in Applied Mechanics and Engineering 197:2408-2430 doi:http://dx.doi.org/10.1016/j.cma.2007.07.030

Fonseca J, Mares C, Friswell M, Mottershead J (2002) Review of parameter uncertainty propagation methods in structural dynamic analysis Proc of ISMA27, Leuven, Belgium:1853-1860

Ghanem RG, Doostan A, Red-Horse J (2008) A probabilistic construction of model validation Computer Methods in Applied Mechanics and Engineering 197:2585-2595

Giunta A, Eldred M Implementation of a trust region model management strategy in the DAKOTA optimization toolkit. In: 8th Symposium on Multidisciplinary Analysis and Optimization, 2000. p 4935

Goodfellow I, Bengio Y, Courville A, Bengio Y (2016) Deep learning vol 1. MIT press Cambridge,

Gorissen D, Couckuyt I, Demeester P, Dhaene T, Crombecq K (2010) A surrogate modeling and adaptive sampling toolbox for computer based design Journal of Machine Learning Research 11:2051-2055

Guo J, Du X (2007) Sensitivity analysis with mixture of epistemic and aleatory uncertainties AIAA journal 45:2337-2349

Haldar A, Mahadevan S (2000) Probability, reliability, and statistical methods in engineering design. John Wiley,

Halder A, Bhattacharya R Model validation: A probabilistic formulation. In: Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on, 2011. IEEE, pp 1692-1697

Harmel RD, Smith PK, Migliaccio KW (2010) Modifying goodness-of-fit indicators to incorporate both measurement and model uncertainty in model calibration and validation Transactions of the ASABE 53:55-63

Hein PRG, Brancheriau L (2018) Comparison between three-point and four-point flexural tests to determine wood strength of Eucalyptus specimens Maderas Ciencia y tecnología 20:333-342

Helton JC, Davis FJ (2003) Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems Reliability Engineering & System Safety 81:23-69

Helton JC, Johnson JD, Sallaberry CJ, Storlie CB (2006) Survey of sampling-based methods for uncertainty and sensitivity analysis Reliability Engineering & System Safety 91:1175-1209

Higdon D, Nakhleh C, Gattiker J, Williams B (2008) A Bayesian calibration approach to the thermal problem Computer Methods in Applied Mechanics and Engineering 197:2431-2441

Hill WJ, Hunter WG (1966) A review of response surface methodology: a literature survey Technometrics 8:571-590

Hills RG (2006) Model Validation: Model Parameter and Measurement Uncertainty Journal of Heat Transfer 128:339 doi:10.1115/1.2164849

Hills RG, Pilch M, Dowding KJ, Red-Horse J, Paez TL, Babuška I, Tempone R (2008) Validation Challenge Workshop Computer Methods in Applied Mechanics and Engineering 197:2375-2380 doi:10.1016/j.cma.2007.10.016

Hu C, Youn BD, Wang P (2018) Engineering Design under Uncertainty and Health Prognostics. Springer,

Ichikawa M (1993) A meaning of the overlapped area under probability density curves of stress and strength Reliability Engineering & System Safety 41:203-204

Jeon B (2016) Unsupervised Learning and Diagnosis Method for Journal Bearing

System in a Large-scale Power Plant. Seoul National University

Jiang X, Mahadevan S (2007) Bayesian risk-based decision method for model validation under uncertainty Reliability Engineering & System Safety 92:707-718

Jiang X, Mahadevan S (2008) Bayesian validation assessment of multivariate computational models Journal of Applied Statistics 35:49-65

Jiang X, Mahadevan S (2011) Wavelet spectrum analysis approach to model validation of dynamic systems Mechanical Systems and Signal Processing 25:575-590

Jung B (2011) A hierarchical framework for statistical model validation of engineered systems.

Jung BC, Lee D, Youn BD, Lee S (2011) A statistical characterization method for damping material properties and its application to structural-acoustic system design Journal of mechanical science and technology 25:1893-1904

Jung BC, Park J, Oh H, Kim J, Youn BD (2014) A framework of model validation and virtual product qualification with limited experimental data based on statistical inference Structural and Multidisciplinary Optimization:1-11

Jung BC, Yoon H, Oh H, Lee G, Yoo M, Youn BD, Huh YC (2016) Hierarchical model calibration for designing piezoelectric energy harvester in the presence of variability in material properties and geometry Structural and Multidisciplinary Optimization 53:161-173

Kang S-C, Koh H-M, Choo JF (2010) An efficient response surface method using moving least squares approximation for structural reliability analysis Probabilistic Engineering Mechanics 25:365-371

Kang Y-J, Lim O-K, Noh Y (2016) Sequential statistical modeling method for distribution type identification Structural and Multidisciplinary Optimization 54:1587-1607

Kang Y-J, Noh Y, Lim O-K (2018) Kernel density estimation with bounded data Structural and Multidisciplinary Optimization 57:95-113

Kat C-J, Els PS (2012) Validation metric based on relative error Mathematical and Computer Modelling of Dynamical Systems 18:487-520

Katayama M (1999) Tft-lcd technology Thin Solid Films 341:140-147

Kennedy MC, O'Hagan A (2001) Bayesian calibration of computer models Journal of the Royal Statistical Society Series B, Statistical Methodology:425-464

Kim B et al. 3.4: Invited Paper: Developments of Transmissive a-Si TFT-LCD using Low Temperature Processes on Plastic Substrate. In: SID Symposium Digest of Technical Papers, 2004. vol 1. Wiley Online Library, pp 19-21

Kleijnen JP (2009) Kriging metamodeling in simulation: A review European Journal of Operational Research 192:707-716

Kokkolaras M, Hulbert G, Papalambros P, Mourelatos Z, Yang R, Brudnak M, Gorsich D (2013) Towards a comprehensive framework for simulation-based design validation of vehicle systems International Journal of Vehicle Design 61:233-248

Kutluay E, Winner H (2014) Validation of vehicle dynamics simulation models – a review Vehicle System Dynamics 52:186-200 doi:10.1080/00423114.2013.868500

Kwaśniewski L (2009) On practical problems with verification and validation of computational models Archives of Civil Engineering 55:323-346

Lee G, Yi G, Youn BD (2018) Special issue: a comprehensive study on enhanced optimization-based model calibration using gradient information Structural and Multidisciplinary Optimization:1-21

Lee I, Choi K, Du L, Gorsich D (2008a) Inverse analysis method using MPP-based dimension reduction for reliability-based design optimization of nonlinear and multi-dimensional systems Computer Methods in Applied Mechanics and Engineering 198:14-27

Lee I, Choi K, Gorsich D (2010) System reliability-based design optimization using the MPP-based dimension reduction method Structural and Multidisciplinary Optimization 41:823-839

Lee I, Choi KK, Du L, Gorsich D (2008b) Dimension reduction method for reliability-based robust design optimization Computers & Structures 86:1550-1562 doi:10.1016/j.compstruc.2007.05.020

Lee I, Choi KK, Du L, Gorsich D (2008c) Inverse analysis method using MPP-based dimension reduction for reliability-based design optimization of nonlinear and multi-dimensional systems Computer Methods in Applied Mechanics and Engineering 198:14-27 doi:10.1016/j.cma.2008.03.004

Lee JH, Gard K (2014) Vehicle–soil interaction: Testing, modeling, calibration and validation Journal of Terramechanics 52:9-21 doi:10.1016/j.jterra.2013.12.001

Lee SH, Chen W (2009) A comparative study of uncertainty propagation methods for black-box-type problems Structural and Multidisciplinary Optimization 37:239-253

Li C, Mahadevan S (2016) Role of calibration, validation, and relevance in multi-level uncertainty integration Reliability Engineering & System Safety 148:32-43 doi:10.1016/j.ress.2015.11.013

Li J, Mourelatos ZP, Kokkolaras M, Papalambros PY, Gorsich DJ Validating designs through sequential simulation-based optimization. In: ASME 2010 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, 2010. American Society of Mechanical Engineers, pp 1023-1031

Li K-C, Aragon Y, Shedden K, Thomas Agnan C (2003) Dimension reduction for multivariate response data Journal of the American Statistical Association 98:99-109

Li W, Chen W, Jiang Z, Lu Z, Liu Y (2014) New validation metrics for models with multiple correlated responses Reliability Engineering & System Safety 127:1-11 doi:10.1016/j.ress.2014.02.002

Liang B, Mahadevan S (2011) Error and uncertainty quantification and sensitivity analysis in mechanics computational models International Journal for Uncertainty Quantification 1

Liang J, Mourelatos ZP, Tu J (2008) A single-loop method for reliability-based design optimisation International Journal of Product Development 5:76-92

Lin G, Elizondo M, Lu S, Wan X (2014) Uncertainty Quantification in Dynamic Simulations of Large-scale Power System Models using the High-Order Probabilistic Collocation Method on Sparse Grids International Journal for Uncertainty Quantification 4

Lin MH, Hsieh CH, Huang KF 72.1: Strength of LCD Panel During Bending. In: SID Symposium Digest of Technical Papers, 2010. vol 1. Wiley Online Library, pp 1069-1072

Ling Y, Mahadevan S (2013) Quantitative model validation techniques: New insights Reliability Engineering & System Safety 111:217-231

Liu H, Chen W, Kokkolaras M, Papalambros PY, Kim HM (2006) Probabilistic analytical target cascading: a moment matching formulation for multilevel optimization under uncertainty Journal of Mechanical Design 128:991-1000

Liu Y, Chen W, Arendt P, Huang H-Z (2011) Toward a better understanding of model validation metrics Journal of Mechanical Design 133:071005

Mahadevan S, Rebba R (2005) Validation of reliability computational models using Bayes networks Reliability Engineering & System Safety 87:223-232

Massey Jr FJ (1951) The Kolmogorov-Smirnov test for goodness of fit Journal of the American statistical Association 46:68-78

McFarland J, Mahadevan S (2008a) Error and variability characterization in structural dynamics modeling Computer Methods in Applied Mechanics and Engineering 197:2621-2631 doi:10.1016/j.cma.2007.07.029

McFarland J, Mahadevan S (2008b) Multivariate significance testing and model calibration under uncertainty Computer methods in applied mechanics and engineering 197:2467-2479

McLachlan G, Krishnan T (2007) The EM algorithm and extensions vol 382. John Wiley & Sons,

Molina A, Al-Ashaab AH, Ellis TI, Young RI, Bell R (1995) A review of computer-aided simultaneous engineering systems Research in Engineering Design 7:38-63

Montgomery DC (2008) Design and analysis of experiments. John Wiley & Sons,

Mousaviraad SM, He W, Diez M, Stern F (2013) Framework for convergence and validation of stochastic uncertainty quantification and relationship to deterministic verification and validation International Journal for Uncertainty Quantification 3

Mullins J, Ling Y, Mahadevan S, Sun L, Strachan A (2016) Separation of aleatory and epistemic uncertainty in probabilistic model validation Reliability Engineering & System Safety 147:49-59 doi:10.1016/j.ress.2015.10.003

Myers RH, Montgomery DC, Anderson-Cook CM (2016) Response surface methodology: process and product optimization using designed experiments. John Wiley & Sons,

Myung IJ (2003) Tutorial on maximum likelihood estimation Journal of mathematical Psychology 47:90-100

Newey WK, West KD (1987) Hypothesis testing with efficient method of moments estimation International Economic Review:777-787

Nguyen TH, Song J, Paulino GH (2010) Single-loop system reliability-based design optimization using matrix-based system reliability method: theory and applications Journal of Mechanical Design 132:011005

Nikolaidis E, Burdisso R (1988) Reliability based optimization: a safety index approach Computers & structures 28:781-788

Oberkampf WL, Barone MF (2006) Measures of agreement between computation and experiment: Validation metrics Journal of Computational Physics 217:5-36 doi:10.1016/j.jcp.2006.03.037

Oberkampf WL, DeLand SM, Rutherford BM, Diegert KV, Alvin KF (2002) Error and uncertainty in modeling and simulation Reliability Engineering & System Safety 75:333-357

Oberkampf WL, Helton JC, Joslyn CA, Wojtkiewicz SF, Ferson S (2004a) Challenge problems: uncertainty in system response given uncertain parameters Reliability Engineering & System Safety 85:11-19

Oberkampf WL, Roy CJ (2010) Verification and validation in scientific computing. Cambridge University Press, Oberkampf WL, Trucano TG (2002) Verification and validation in computational fluid dynamics Progress in Aerospace Sciences 38:209-272

Oberkampf WL, Trucano TG (2008) Verification and validation benchmarks Nuclear Engineering and Design 238:716-743 doi:10.1016/j.nucengdes.2007.02.032

Oberkampf WL, Trucano TG, Hirsch C (2004b) Verification, validation, and predictive capability in computational engineering and physics Applied Mechanics Reviews 57:345-384

Oden JT, Prudencio EE, Bauman PT (2013) Virtual model validation of complex multiscale systems: Applications to nonlinear elastostatics Computer Methods in Applied Mechanics and Engineering 266:162-184 doi:10.1016/j.cma.2013.07.011

Oh H, Choi H, Youn BD, Jung JH A Robust Metric in Optimization-Based Approach for Statistical Model Calibration. In: presented at the 12th World Congress on Structural and Multidisciplinary Optimization, Braunschweig, Germany, June 5-9 2017.

Oh H, Kim J, Son H, Youn BD, Jung BC (2016) A systematic approach for model refinement considering blind and recognized uncertainties in engineered product development Structural and Multidisciplinary Optimization 54:1527-1541

Oliver TA, Terejanu G, Simmons CS, Moser RD (2015) Validating predictions of unobserved quantities Computer Methods in Applied Mechanics and Engineering 283:1310-1335 doi:10.1016/j.cma.2014.08.023

Pak CH, Lee U-S, Hong SC, Song SK, Kim JH, Kim KS (1991) A study on the tangential vibration of the steering wheel of passenger car. SAE Technical Paper, Pan C, Hsieh C, Su C, Liu Z (2008) Study of cutting quality for TFT-LCD glass substrate The International Journal of Advanced Manufacturing Technology 39:1071-1079

Pan M-C, Chen P-C (2007) Drop simulation/experimental verification and shock resistance improvement of TFT–LCD monitors Microelectronics Reliability 47:2249-2259

Park I, Amarchinta HK, Grandhi RV (2010) A Bayesian approach for quantification of model uncertainty Reliability Engineering & System Safety 95:777-785 doi:10.1016/j.ress.2010.02.015

Parry GW (1996) The characterization of uncertainty in probabilistic risk assessments of complex systems Reliability Engineering & System Safety 54:119-126

Plackett RL (1983) Karl Pearson and the chi-squared test International Statistical Review/Revue Internationale de Statistique:59-72

Quinn GD, Sparenberg BT, Koshy P, Ives LK, Jahanmir S, Arola DD (2009) Flexural strength of ceramic and glass rods Journal of Testing and Evaluation 37:222-244

Rahman S, Xu H (2004) A univariate dimension-reduction method for multi-dimensional integration in stochastic mechanics Probabilistic Engineering Mechanics 19:393-408

Rebba R, Mahadevan S (2006) Validation of models with multivariate output Reliability Engineering & System Safety 91:861-871

Rebba R, Mahadevan S (2008) Computational methods for model reliability assessment Reliability Engineering & System Safety 93:1197-1207

Roussouly N, Petitjean F, Salaun M (2013) A new adaptive response surface method for reliability analysis Probabilistic Engineering Mechanics 32:103-115

Roy CJ, Oberkampf WL A complete framework for verification, validation, and uncertainty quantification in scientific computing. In: 48th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition, 2010. pp 4-7

Roy CJ, Oberkampf WL (2011) A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing Computer Methods in Applied Mechanics and Engineering 200:2131-2144

Sankararaman S, Ling Y, Mahadevan S (2011) Uncertainty quantification and model validation of fatigue crack growth prediction Engineering Fracture Mechanics 78:1487-1504

Sankararaman S, Mahadevan S (2011) Model validation under epistemic uncertainty Reliability Engineering & System Safety 96:1232-1241

Sankararaman S, Mahadevan S (2015) Integration of Model Verification, Validation, and Calibration for Uncertainty Quantification in Engineering Systems Reliability Engineering & System Safety

Sargent RG (2013) Verification and validation of simulation models Journal of Simulation 7:12-24

Sarin H, Kokkolaras M, Hulbert G, Papalambros P, Barbat S, Yang R-J (2010) Comparing time histories for validation of simulation models: error measures and metrics Journal of dynamic systems, measurement, and control 132:061401

Scholz F (1985) Maximum likelihood estimation Encyclopedia of statistical sciences

Schuëller GI, Jensen HA (2008) Computational methods in optimization considering uncertainties – An overview Computer Methods in Applied Mechanics and

Engineering 198:2-13 doi:10.1016/j.cma.2008.05.004

Schwer LE (2007) Validation metrics for response histories: perspectives and case studies Engineering with Computers 23:295-309

Shi L, Yang R, Zhu P (2012) A method for selecting surrogate models in crashworthiness optimization Structural and Multidisciplinary Optimization 46:159-170

Silva AS, Ghisi E (2014) Uncertainty analysis of the computer model in building performance simulation Energy and Buildings 76:258-269 doi:10.1016/j.enbuild.2014.02.070

Smirnov N (1948) Table for estimating the goodness of fit of empirical distributions The annals of mathematical statistics 19:279-281

Snyman J (2005) Practical mathematical optimization: an introduction to basic optimization theory and classical and new gradient-based algorithms vol 97. Springer Science & Business Media,

Soundappan P, Nikolaidis E, Haftka RT, Grandhi R, Canfield R (2004) Comparison of evidence theory and Bayesian theory for uncertainty modeling Reliability engineering & System safety 85:295-311

Stephens MA (1974) EDF statistics for goodness of fit and some comparisons Journal of the American statistical Association 69:730-737

Su C-T, Hsiao Y-H, Liu Y-L (2012) Enhancing the fracture resistance of medium/small-sized TFT-LCDs using the Six Sigma methodology IEEE Transactions on Components, Packaging and Manufacturing Technology 2:149-164

Swiler LP, Adams BM, Eldred MS (2008) Model calibration under uncertainty: matching distribution information Proceedings of the 12th AIAA/ISSMO

Tekin E, Sabuncuoglu I (2004) Simulation optimization: A comprehensive review on theory and applications IIE Transactions 36:1067-1081 doi:10.1080/07408170490500654

Terzaghi K (1944) Theoretical soil mechanics. Chapman And Hali, Limited John Wiler And Sons, Inc; New York,

Thacker BH, Paez TL A Simple Probabilistic Validation Metric for the Comparison of Uncertain Model and Test Results. In: ASME Verification and Validation Symposium, 2013.

Trucano TG, Swiler LP, Igusa T, Oberkampf WL, Pilch M (2006) Calibration, validation, and sensitivity analysis: What's what Reliability Engineering & System Safety 91:1331-1357 doi:10.1016/j.ress.2005.11.031

Tsai C-H, Lin B-C (2007) Laser cutting with controlled fracture and pre-bending applied to LCD glass separation The International Journal of Advanced Manufacturing Technology 32:1155-1162

Tu J, Choi KK, Park YH (1999) A new study on reliability-based design optimization Journal of mechanical design 121:557-564

Urbina A, Mahadevan S, Paez TL (2011) Quantification of margins and uncertainties of complex systems in the presence of aleatoric and epistemic uncertainty Reliability Engineering & System Safety 96:1114-1125

Uusitalo L, Lehikoinen A, Helle I, Myrberg K (2015) An overview of methods to evaluate uncertainty of deterministic models in decision support Environmental Modelling & Software 63:24-31 doi:10.1016/j.envsoft.2014.09.017

van Beers WCM, Kleijnen JPC (2008) Customized sequential designs for random simulation experiments: Kriging metamodeling and bootstrapping European

Journal of Operational Research 186:1099-1113 doi:10.1016/j.ejor.2007.02.035

Venter G (2010) Review of optimization techniques Encyclopedia of aerospace engineering

Vepakomma KH, Westbrook J, Carley S, Kim J (2013) Finite element analysis of ring-on-ring test on LCD panels Journal of Display Technology 9:673-677

Voyles IT, Roy CJ Evaluation of model validation techniques in the presence of aleatory and epistemic input uncertainties. In: 17th AIAA Non-Deterministic Approaches Conference, ARC, 2015. pp 1-16

Weathers JB, Luck R, Weathers JW (2009) An exercise in model validation: Comparing univariate statistics and Monte Carlo-based multivariate statistics Reliability Engineering & System Safety 94:1695-1702 doi:10.1016/j.ress.2009.04.007

White H (1982) Maximum likelihood estimation of misspecified models Econometrica: Journal of the Econometric Society:1-25

XI Z, FU Y, YANG R-J (2013) Model bias characterization in the design space under uncertainty International Journal of Performability Engineering 9:433

Xie K, Wells L, Camelio JA, Youn BD (2007) Variation propagation analysis on compliant assemblies considering contact interaction Journal of Manufacturing Science and Engineering 129:934-942

Xiong Y, Chen W, Tsui K-L, Apley DW (2009) A better understanding of model updating strategies in validating engineering models Computer Methods in Applied Mechanics and Engineering 198:1327-1337 doi:10.1016/j.cma.2008.11.023

Xu H, Rahman S (2004) A generalized dimension-reduction method for multidimensional integration in stochastic mechanics International Journal for Numerical Methods in Engineering 61:1992-2019

Xu X, Wang L, Newman ST (2011) Computer-aided process planning – A critical review of recent developments and future trends International Journal of Computer Integrated Manufacturing 24:1-31 doi:10.1080/0951192x.2010.518632

Yao W, Chen X, Luo W, van Tooren M, Guo J (2011) Review of uncertainty-based multidisciplinary design optimization methods for aerospace vehicles Progress in Aerospace Sciences 47:450-479

Yates F (1934) Contingency tables involving small numbers and the χ 2 test Supplement to the Journal of the Royal Statistical Society 1:217-235

Yoon J (2018) Resilience Analysis and Design Methodology Considering False Alarms for Complex Engineered Systems. Seoul National University

Youn BD, Choi K, Yang R-J, Gu L (2004) Reliability-based design optimization for crashworthiness of vehicle side impact Structural and Multidisciplinary Optimization 26:272-283

Youn BD, Jung BC, Xi Z, Kim SB, Lee W (2011) A hierarchical framework for statistical model calibration in engineering product development Computer Methods in Applied Mechanics and Engineering 200:1421-1431

Youn BD, Wang P (2008) Bayesian reliability-based design optimization using eigenvector dimension reduction (EDR) method Structural and Multidisciplinary Optimization 36:107-123

Youn BD, Xi Z (2009) Reliability-based robust design optimization using the eigenvector dimension reduction (EDR) method Structural and Multidisciplinary Optimization 37:475-492

Youn BD, Xi Z, Wang P (2008) Eigenvector dimension reduction (EDR) method for sensitivity-free probability analysis Structural and Multidisciplinary Optimization 37:13-28

Yuan J, Ng SH (2013) A sequential approach for stochastic computer model calibration and prediction Reliability Engineering & System Safety 111:273-286

Zaremba A, Liubakka M, Stuntz R Control and steering feel issues in the design of an electric power steering system. In: American Control Conference, 1998. Proceedings of the 1998, 1998. IEEE, pp 36-40

Zhan Z, Fu Y, Yang R-J, Peng Y (2011) An automatic model calibration method for occupant restraint systems Structural and Multidisciplinary Optimization 44:815-822

Zhang E, Feissel P, Antoni J (2011) A comprehensive Bayesian approach for model updating and quantification of modeling errors Probabilistic engineering mechanics 26:550-560

Zhang R, Mahadevan S (2000) Model uncertainty and Bayesian updating in reliability-based inspection Structural Safety 22:145-160

Zhang S, Zhu P, Chen W, Arendt P (2013) Concurrent treatment of parametric uncertainty and metamodeling uncertainty in robust design Structural and Multidisciplinary Optimization 47:63-76

Zhou Z, Ong Y-S, Nair PB, Keane AJ, Lum KY (2007) Combining global and local surrogate models to accelerate evolutionary optimization IEEE Transactions On Systems, Man and Cybernetics-Part C 37:66-76

Zhu S-P, Huang H-Z, Peng W, Wang H-K, Mahadevan S (2016) Probabilistic Physics of Failure-based framework for fatigue life prediction of aircraft gas turbine discs under uncertainty Reliability Engineering & System Safety 146:1-12 doi:10.1016/j.ress.2015.10.002

Acknowledgement

First, I would like to express my sincere gratitude to my advisor Prof. Byeng D. Youn for the continuous support of my Ph.D study. His guidance patiently helped me through all the time of research and the dissertation process, continuously challenging me to set my sight higher and never accepting less than my best efforts. I could not have imagined having a better advisor and mentor for my Ph.D study.

Besides my advisor, I would like to appreciate the rest of my thesis committee: Prof. Yoon Young Kim, Prof. Maenghyo Cho, Prof. Do-Nyun Kim, and Prof. Hyunseok Oh, for their insightful comments and encouragement, but also for the valuable questions, which incented me to widen my research from various perspectives.

My sincere thanks also goes to Dr. Guilian Yi, Dr. Taejin Kim, and Dr. Heonjun Yoon for the stimulating discussions related to my research. Without their precious support, it would not be possible to develop my Ph.D research.

I thank my labmates including Ph.D alumni (Dr. Byung Chul Jeon, Dr. Jong Moon Ha, Dr. Dae Whan Kim, Dr. Joung Taek Yoon, Dr. Woosung Choi), M.Eng. alumni (Chulmin Cho, Kyung min Park, Jinshi Cui, Sihyeong Woo, Hwanoh Choi, Keon Kim, Byungjoo Jeon, Dongki Lee), and colleagues studying together (Hyunjae Kim, Jungho Park, Keunsu Kim, Joon Ha Jung, Boseong Seo, Minji Yoo, Jin-Gyun Park, Junmin Lee, Taewan Hwang, Beom Chan Jang, Yong Chang Shin, Sunuwe Kim, Su Jii Kim, Myungyon Kim, Chan Hee Park, Soo-Ho Jo, Yunhan Kim, Kyunmin Na, Jin Uk Ko, Jongmin Park, Sooho Kim, Jingyo Heong, Jongkeun Lim, Joowhan Song, Hao Yuan, Gil Jun Ahn, Hyunhee Choi, Hwayong Jeong, Hyeon Bae Kong, Myeongbaek Youn, Yeong Tak Oh, Joohyeong Im). I would like to express my special thanks to Hyejeong Son and Wongon Kim for sharing innumerous days and nights in the lab and their encouragements and support whenever I was in need.

Last but not the least, I wish to extend my special and sincere thanks to my family: my parents, Changhwan Lee & Myungsin Maeng, and to my younger brother, Guesang Lee, for their love and unwavering support during my study as well as my life in general. Thank you to my beloved, Songhyeun Kim, for all her love and support.

국문 초록

최적화 기반 모델 보정을 위한 순차적 최적화 및 불확실성 확산 기법

이 규 석

기계항공공학부 기계공학과

공과대학

서울대학교

높은 신뢰도를 만족하는 전산 모델을 구축하는 것은 오랜 기간 엔지니어들의 꿈이었다. 전산 모델을 통한 높은 신뢰도의 예측 및 해석을 위해선 정확한 입력 변수 값이 필요하다. 하지만, 미지 입력 변수의 존재는 전산 모델을 이용한 예측 및 해석의 신뢰도를 저하하는 대표적인 이유 중 하나이다. 모델 보정이란, 전산 모델의 미지 입력 변수의 값을 추정하는 과정을 뜻하며, 궁극적으로 정확한 미지 입력 변수 값의 추정을 통한 전산 모델의 신뢰도 향상을 목표로 한다. 높은 신뢰도의 전산 모델을 구축하기 위한 많은 관심과 함께, 모델 보정 기법 개발을 위한 많은 연구들이 진행되었으며, 대표적인 방법이 최적화 기반 모델 보정이다.

올바른 최적화 기반 모델 보정을 위해선 다양한 종류의 불확실성 요인들에 대한 고려, 동시에 효율적이고 정확한 모델 보정 방법 개발이 필요하다. 본 논문은 최적화 기반 모델 보정을 이용한 통계적 모델 보정 기법 개발을 위하여 다음 네 가지의 기술적인 문제를 해결하고자 한다. 연구주제 1: 측정 오류를 고려한 실험 데이터의 불확실성 모델링, 2: 최적화 기반 모델 보정의 볼록 최적화 문제 확인을 위한 민감도 정보 유도 및 민감도 정보를 이용한 강건 최적화 기반 모델 보정, 3: 정확하고 효율적인 순차적 통계 기반 최적 설계 루프 구성, 4: 최적화 기반 모델 보정 결과에 대한 직관적이고 통계적인 검증 방법 개발

연구주제 1: 모델 보정은 전산 모델을 이용한 해석 값과 실험 값으로부터 얻은 두 확률 분포 일치도의 최대화(혹은 불일치도의 최소화)를 통해 미지 입력 변수 값을 추정한다. 정확한 모델 보정을 위해선 실험 측정 값의 불확실성을 적절히 묘사하는 확률 분포가 필요하다. 하지만, 실험 중에 발생하는 시스템 측정 오류와 랜덤 측정 오류에 의해 실험 값의 불확실성이 부정확하게 모델링 될 수 있다. 따라서, 첫번째 연구 주제는 최대 우도 추정 방법을 이용하여, 실험 측정 시 발생하는 측정 오류들을 고려한 실험 데이터의 불확실성 모델링을 목표로 한다.

연구주제 2: 최적화 기반 모델 보정은 효율적인 최적화 알고리즘으로 주로 구배 기반 최적화 알고리즘(Gradient-based optimization algorithm)을 사용한다. 하지만, 지금까지 개발된 보정 척도(Calibration metric)를 이용하고, 구배 기반 최적화 알고리즘을 통해 최적화 기반 모델 보정 결과, 부정확하고 불안정한 보정 결과를 보여왔다. 따라서, 두번째 연구 주제는 1) 위 부정확하고 불안정한 보정

결과에 대한 이유를 분석하고, 2) 유도된 민감도 정보를 이용하여 안정적인 최적화 기반 모델 보정 기법을 제안한다.

연구주제 3: 최적화 기반 모델 보정은 통계적인 방법으로 미지 입력 변수의 값을 추정하기 위하여 통계 기반 최적 설계(Optimization under uncertainty)를 사용한다. 통계 기반 최적 설계는 확률 분석(Probabilistic analysis)을 포함한 최적화 방법이다. 최적화 기반 모델 보정을 위해서 확률 분석은 해석 값의 확률 분포를 얻기 위한 불확실성 확산(Uncertainty propagation)을 담당한다. 하지만, 확률 분석은 많은 양의 계산을 필요하므로 최적화 과정과 동시에 진행될 때, 효율적인 최적화 루프를 구성하는 것이 필요하다. 따라서, 세번째 연구는 최적화 기반 모델 보정의 효율성을 높이기 위하여 순차적 최적화 기반 모델 보정 기법을 제안한다. 제안된 순차적 최적화 기반 모델 보정은 두 개의 순차적 최적화 루프로 구성되어 있으며, 첫번째 루프는 효율성을, 두번째 루프는 정확성을 목표로 한다. 제안된 방법을 순차적 최적화 및 불확실성 확산(Sequential optimization and uncertainty propagation: SOUP)이라 한다.

연구주제 4: 모델 보정의 마지막 과정으로, 모델 검증 과정은 보정된 결과가 유효한 결과인지 확인한다. 검증은 정량적이고 통계적인 방법으로 수행되어야 한다. 네번째 연구주제는 새로운 검증 척도인 확률 일치도(Probability of coincidence: POC)를 제안한다. 확률 일치도는 전산 모델로부터 얻은 해석 결과와 실험으로부터 얻은 측정 결과의 확률적인 일치도를 계산한다.

주요어: 최적화 기반 모델 보정
통계적 모델 보정 및 검증
불확실성 모델링
시스템/랜덤 측정 오류
통계 기반 최적화
순차적 최적화 및 불확실성 확산
확률 일치도

학 번: 2014-21865