공학박사학위논문

컴퓨터 모델 내 오류 원인 식별을 위한 최적화 기반 모델 개선 기법 연구

Optimization-based Model Improvement for Error Sources Identification in a Computational Model

2021 년 8 월

서울대학교 대학원

기계항공공학부

손 혜 정

컴퓨터 모델 내 오류 원인 식별을 위한 최적화 기반 모델 개선 기법 연구

Optimization-based Model Improvement for Error Sources Identification in a Computational Model

지도교수 윤 병 동

이 논문을 공학박사 학위논문으로 제출함

2021 년 04 월

서울대학교 대학원
기계항공공학부
손 혜 정

손혜정의 공학박사 학위논문을 인준함

2021 년 06 월

위 원 장 : 김 도 년

부위원장 : 윤 병 동

위 원 : 김 윤 영

위 원 : 안 성 훈

위 원 : 김 태 진

Abstract

Optimization-based Model Improvement for Error Sources Identification in a Computational Model

Hyejeong Son
Department of Mechanical and Aerospace Engineering
The Graduate School
Seoul National University

The increased use of computer-aided engineering (CAE) in recent years requires a more accurate prediction capability in computational models. Therefore, extensive studies have considered engineering strategies to achieve highly credible computational models. Optimization-based model improvement (OBMI), which includes model calibration, validation, and refinement, is one crucial technique that has emerged to enhance the prediction ability of computational models. *Model calibration* is the process of estimating unknown input parameters in a computational model. *Model validation* presents a judgement of the accuracy of a predicted response. If it is possible for a computational model to have model form uncertainties, *model refinement* explores unrecognized error sources of a computational model. OBMI can adopt these three

processes individually or sequentially, according to the trustworthiness of the prior knowledge of the computational modeling.

Although OBMI process improvements have emerged to try to consider the major sources of errors, OBMI can still suffer from a failure to improve a computational model. Since numerous error sources in an experimental and computational model are intertwined with each other, OBMI has difficulty identifying the error sources required to enable accurate prediction ability of the computational model. Thus, eventually, OBMI may fail to propose an appropriate solution. To cope with this challenge, this doctoral dissertation research addresses three essential issues: 1) Research Thrust 1 – a new experimental design approach for model calibration to reduce parameter estimation errors; 2) Research Thrust 2) – a device bias quantification method for considering model form errors with bound information; and, Research Thrust 3) – comparison of statistical validation metrics to consider type II errors in model validation.

Research Thrust 1: A variety of sources of errors in observation and prediction can interrupt the model improvement process. These error sources degrade the parameter estimation accuracy of the model calibration. When a computational model turns out to be invalid because of these error sources, the OBMC process performs model refinement. However, since model validation cannot distinguish between parameter estimation errors and modeling errors, it is difficult for the existing method to efficiently refine the computational model. Thus, this study aims to develop a model improvement process that identifies the leading cause of invalidity of a prediction. In this work, an experimental design method is integrated with optimization-based model improvement to minimize the effect of estimation errors in model calibration. Through

use of the proposed method, after calibration, the computational model mainly includes the effects of unrecognized modeling errors.

Research Thrust 2: The experimental design method proposed in Research Thrust 1 has the advantage of being able to identify two error sources without additional observation. However, model calibration still suffers from parameter estimation errors, since experimental design is affected by model form errors. The parameters estimated by model calibration are often unreasonable for engineers in practical settings because they have expert-based prior knowledge about the model parameters. Among the variety of physical information available, bound information about model parameters is a suitable constraint in optimization-based model calibration (OBMC). Using prior information about parameter bounds, Research Thrust 2 devises proportionate bias calibration to quantify the amount of degradation of the predicted responses that is due to model form errors in a computational model. The bias term is estimated in the optimization-based model calibration (OBMC) algorithm with unknown parameters to enable OBMC to support accurate estimation of unknown parameters within a prior bound. This study proposes a new formulation of a bias term that depends on the output responses to resolve the gap in appropriate bias that arises due to the different dimensions of the predicted responses.

Research Thrust 3: Statistical model validation (SMV) evaluates the accuracy of a computational model's predictions. In SMV, hypothesis testing is used to determine the validity or invalidity of a prediction, based on the value of a statistical validation metric that quantifies the difference between the predicted and observed results. Errors in hypothesis testing decisions are troublesome when evaluating the accuracy of a computational model, since an invalid model might be used in practical engineering

design activities and incorrect results in these settings may lead to safety issues. This research compares various statistical validation metrics to highlight those that show fewer errors in hypothesis testing. The resulting work provides a statistical validation metric that is sensitive to a discrepancy in the mean or variance of the two distributions from the predictions and observations. Statistical validation metrics examined in this study include Kullback-Leibler divergence, area metric with U-pooling, Bayes factor, likelihood, probability of separation, and the probability residual.

Keywords: Optimization-based Model Improvement (OBMI)
Error Source Identification
D-optimality based Experimental Design
Proportionate Bias Calibration
Statistical Validation Metric
Hypothesis Testing
Digital Twin

Student Number: 2015-20733

Table of Contents

List of Tables

List of Figures

Nomenclatures

CAE	Computer-aided engineering
OBMI	Optimization-based model improvement
OBMC	Optimization-based model calibration
SMV	Statistical model validation
ICT	Information and Communication Technology
IoT	Internet of Things
AI	Artificial Intelligence
CPS	Cyber-physical systems
CAM	Computer-aided modeling
LB	Lower bound
UB	Upper bound
KLD	Kull-back Leibler Divergence
PoS	Probability of Separation
PR	Probability Residual
S-Type	Smaller-the-better metrics
L-Type	Larger-the-better metrics
RMSE	Root mean squared errors
α	Confidence level
p	probability
S.t.d.	Standard deviation
PDF	Probability of density function
CDF	Cumulative density function
max	maximization
min	minimization

lim	limiting behaviors
^{T}	Transpose matrix
$\nabla_{\boldsymbol{\theta}}$	gradient operator about $\boldsymbol{\theta}$
$^{-1}$	Inverse matrix
$\lvert \ \rvert$	Absolute value
det	Determinant
$\sum$	summation
$\prod$	multiplication
ln	Natural logarithm
$\int$	integral
$\frac{\partial}{\partial E'}, \frac{\partial}{\partial W}$	Partial derivative of the parameters E and W
∞	infinite
$\boldsymbol{\theta}, \boldsymbol{\theta}_{\alpha}$	Unknown model parameter vector
$\boldsymbol{\theta}_{B_k}$	Unknown model parameter vector whose elements consists of lower bound or upper bound of each parameters
$\boldsymbol{x}$	Controllable design parameter vector
$\mathbf{D}$	Experimental design vector
f_{cal}	Function of calibration metric
$\boldsymbol{\eta}$	Predicted response
$\boldsymbol{\eta}_{\boldsymbol{true}}$	Predicted response without model form error
$\boldsymbol{\zeta}$	Observed (observation) response
p	Probability function
ζ_i	Observation data
$\mathbf{N\sim()}$	Normal distribution
$\boldsymbol{\varepsilon}$	Experimental error, measurement error
$\boldsymbol{\delta}$	Model form error vector
δ_{bias}	Model bias term representing an assumed model form error

(m×n)	Matrix size
$\boldsymbol{\theta_e}$	Estimates of unknown parameters vector
$\mathbf{D_c}$	Experimental design vector for model calibration
$\mathbf{D_v}$	Experimental design vector for model validation
P	Input point load for a cantilever beam
H	height of a cantilever beam
W	width of a cantilever beam
L	Length of a cantilever beam
σ_{true}	Stress of a cantilever beam
$\sigma_{\boldsymbol{\eta}}$	Predicted stress of a cantilever beam
$\sigma_{\boldsymbol{\zeta}}$	Observed stress of a cantilever beam
u_{true}	Displacement of a cantilever beam
$u_{\boldsymbol{\eta}}$	Predicted displacement of a cantilever beam
$u_{\boldsymbol{\zeta}}$	Observed displacement of a cantilever beam
E	Young's modulus
$d_{c,\sigma}$	Experimental design of the stress for model calibration
$d_{c,u}$	Experimental design for the displacement for model calibration
W_e, E_e	Estimates of width and Young's modulus
L_δ	Length of a cantilever beam in a prediction model due to model form error
$\mathbf{M}$	Design selection metric for experimental design
$\varepsilon_\sigma, \varepsilon_u.$	Measurement error for stress and displacement
$\boldsymbol{\psi_\eta}$	Prediction errors
ψ_σ	Prediction error of the stress at all of measured location
ψ_u	Prediction error of the displacement at all of measured location
[,]	Range of a specific parameter
ρ	Poission's ratio

Pr	Pressure of a tire
Pr_e	Estimate for pressure of a tire
C	Combination behavior
$\mathbf{L}_\delta$	Lower bound of the model bias
$\mathbf{U}_\delta$	Upper bound of the model bias
$\varepsilon rr^*_{\theta_i}$	Parameter estimation error of unknown model parameters
C_1,C_2	Distribution of the prediction or observation data
$\tilde{x}_{c_i}$	median
f	Probability density function
F	Cumulative density function
H_0	Null hypothesis
H_1	Alternative hypothesis
A	Area metric
A_u	Area metric with U-pooling method
u_i	U-value
Lh	Likelihood
D_{KL}	Kullback-Liebler divergence
B_0	Bayes factor
f^{prr}	Prior PDF
θ_{pred}	A value of unknown parameters in initial prediction
θ_{obs}	A value of unknown parameters in observation
S	Scale parameter for PR
G	A nonlinear function for Chapter 5
n	The number of observation datasets evaluated correct
N	The number of observation datasets
Φ	CDF of standard normal distribution

$\mu_{1,}\mu_2$	First moment
σ_1,σ_2	Second moment
m_1, m_2	mean
s_1, s_2	Standard deviation

Chapter 1

Introduction

1.1 Motivation

Advances in Information and Communication Technology (ICT) have triggered an Industry 4.0 trend in manufacturing systems. Based on state-of-art ICT technologies, such as Internet of Things (IoT) platforms, Artificial Intelligence (AI), and cyber-physical-systems (CPS), Industry 4.0 has provided automatic operation in design, management, and monitoring of the entire manufacturing process (Peraković et al. 2018) (Alcácer and Cruz-Machado 2019) (Peraković et al. 2019). The recent trend of Industry 4.0-based manufacturing has given computer-aided modeling (CAM) a crucial role in the product development and management processes (Mosterman and Zander 2016) (Qi and Tao 2018). CAM has enabled more accurate prediction ability in a wide range of manufacturing settings. Furthermore, the most recent types of CAM have evolved to enable digital twin exploration of cyber-physical systems (CPS) that can predict responses in diverse operating conditions (e.g., system deformation or environmental change), while considering numerous model parameters (Mosterman and Zander 2016). However, some digital twin parameters become invalid as the system condition changes; this decreases the prediction accuracy of this approach. Thus, model improvement techniques have been

developed to enhance the capabilities of computational models. (Kennedy and O'Hagan 2001) (Trucano et al. 2006) (Xiong et al. 2009) (Oberkampf and Roy 2010) (Youn et al. 2011) (Ling et al. 2013) (Sun and Sun 2015) (Moon et al. 2015) (Lee et al. 2018) (Lee et al. 2019) (Hu et al. 2020).

Discovering unknown input parameters in a computational model is one of the core tasks of model improvement. Optimization-based model calibration (OBMC) is recognized as a promising solution for estimating unknown input parameters in a computational model, through the use of optimization techniques. Optimization-based model improvement (OBMI) is terminology for the model improvement process, which employs OBMC to estimate unknown input parameters. Including the model calibration step, OBMI consists of three sub-steps; model calibration, model validation, and model refinement. Model validation is a decision-making step to evaluate the prediction accuracy of a computational model based upon observation data. If the model validation step decides that a computational model is invalid, the process of OBMI assumes the possibility of blind sources of model form errors. Then, model refinement examines the most dominant error sources that were unrecognized in the modeling, through a systematic framework. The overall flow chart of OBMI and its three sub-steps are shown in Figure 1-1 (Youn et al. 2011).

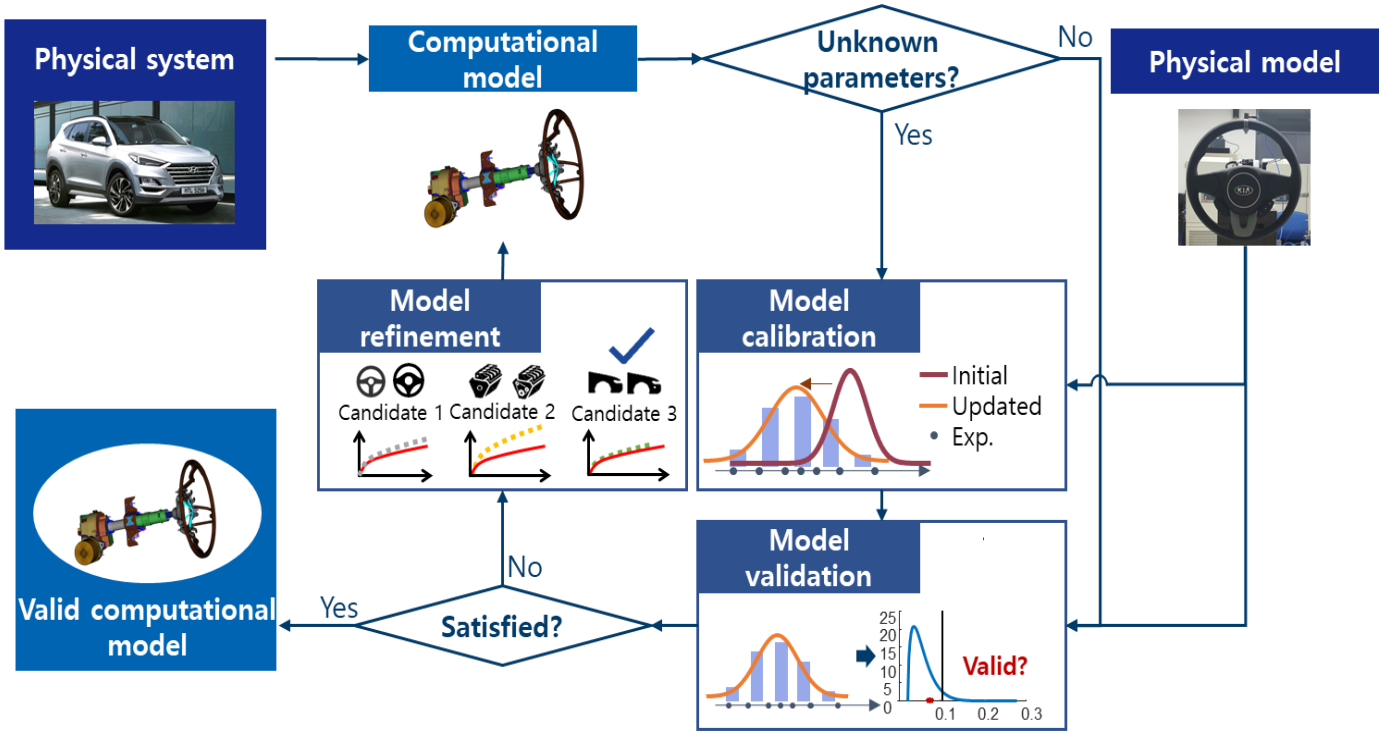


Figure 1-1 A flow chart of Optimization-based Model Improvement (OBMI) for a computational model

Despite the extensive studies about model improvement, engineers who use the OBMI process still face limitations in their ability to enhance the predictability of a computational model. A variety of error sources in the computational model and observation data can threaten the ability to achieve a feasible OBMI result. Figure 1-2, Figure 1-3, and Figure 1-4 illustrate three major error sources that directly affect observation data and computational models. The model parameters include material properties, boundary and initial conditions, geometric conditions, and environmental conditions, as shown Figure 1-2. Model form errors are a comprehensive term for modeling failures that arise due to insufficient knowledge, such as model simplification, surrogate modeling errors, and wrong assumptions. The last error sources are experimental errors, as illustrated in Figure 1-4. Since the predicted responses of a computational model are not directly influenced by experimental errors, this is regarded as the least influential factor. However, experimental errors should be considered, because the observation data serves as a criterion for estimating parameters in the model calibration and for evaluating the validity of a computational model in the model validation step.

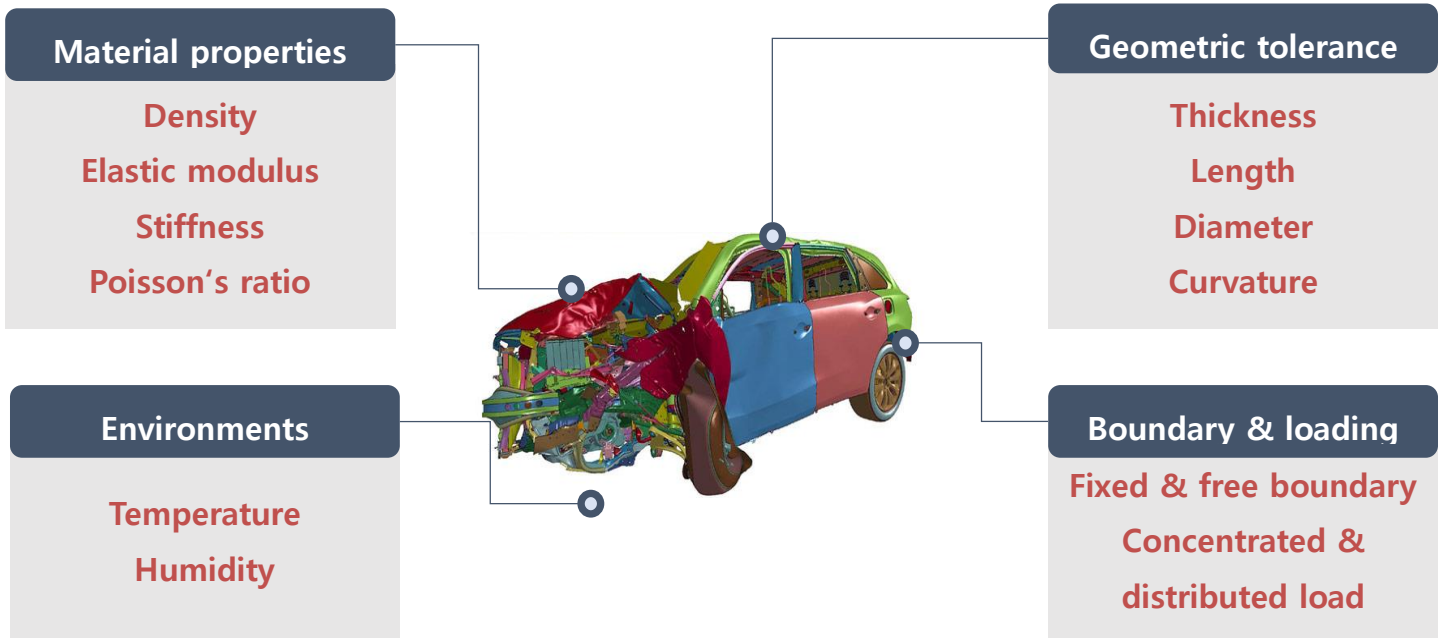
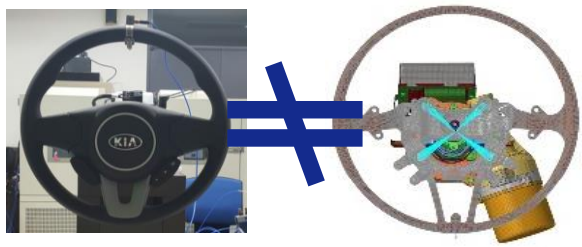
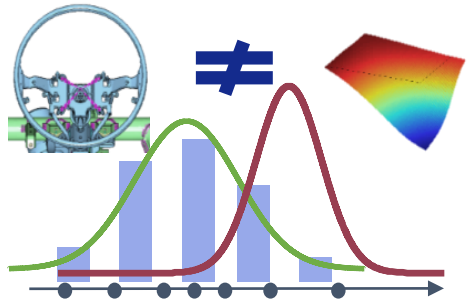


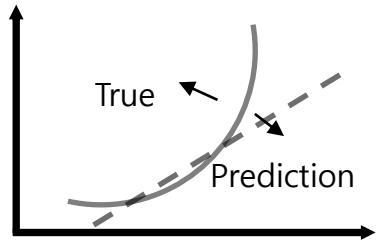
Figure 1-2 Various error sources in OBMI: Unknown model parameters



(a)



(b)



(c)

Figure 1-3 Various error sources in OBMI: Model form errors: (a) simplified modeling, (b) surrogate modeling, and (c) wrong assumptions

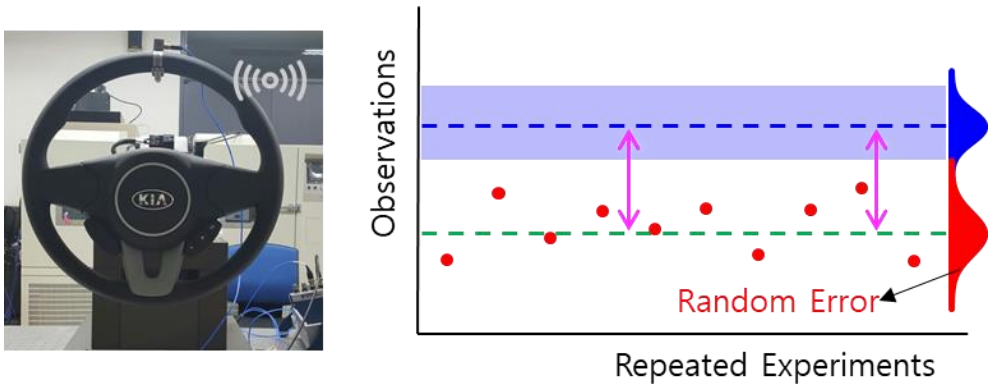


Figure 1-4 Various error sources in OBMI: Experimental errors

Due to these error sources, model improvement remains a challenging task for development of an accurate process. The first challenge is that the estimated parameters used in model calibration cannot ensure that the result is similar to the actual value of a real system (Oberkampf and Trucano 2008) (Oberkampf and Roy 2010). The estimated parameter values in model calibration allow the computational model to emulate the observation data used for model calibration. It is unknown whether the estimated parameter values are similar to the values of the actual system. In the end, model improvement cannot guarantee the validity of the estimated parameter values in other designs without new observation data for the new design. The second challenge is that model validation cannot ensure an accurate decision. An invalid computational model can be evaluated as a valid model due to the various sorts of uncertainties in the predicted responses and observation data. These decision errors can be dangerous in actual engineering fields, especially when designing a vast and complex engineering system related to human safety. A trustworthy result of OBMI can be available after these challenges are solved. Thus, section 1.2 outlines the three technical issues that should be solved to enable wide use of OBMI,

through consideration of major error sources.

1.2 Research Scope and Overview

An accurate OBMI process must be performed with reasonable consideration of the various error sources. The goal of this doctoral dissertation research is to develop an advanced OBMI process to enable wide use of a computational model, by tackling three technical issues: 1) the limitations of OBMI that arise due to coupled error sources in the model calibration; 2) parameter estimation errors in model calibration that emerge due to the existence of unrecognized model form errors; 3) decision errors in statistical model validation that are due to numerous sources of uncertainties. To address the above-mentioned technical issues, the research scope in this doctoral dissertation is the enhancement of OBMI through the following three research thrusts:

Research Thrust 1: Experimental Design to Identify Error Sources in Optimization-based Model Improvement

Due to the error sources introduced in Fig 2-4, it is difficult for the model calibration and model validation steps to offer reliable operation. One of the main problems in model calibration is that the error sources degrade the estimation accuracy of the unknown model parameters. These parameter estimation errors in model calibration become another error source that is coupled with model form errors in the computational model. Since OBMC cannot discover the existence of these parameter estimation errors, a challenge in OBMI arises in that the process cannot identify the

coupled error sources in the model validation (Figure 1-5). These unidentified errors can prevent model validation from giving the correct decision. The objective of Research Thrust 1 is to identify two sorts of error sources between parameter estimation errors and model form errors to enable accurate decisions in model validation. The proposed method introduces an experimental design approach for model calibration that reduces the parameter estimation errors to address this issue. The optimality criterion for the experimental design is derived from the analytical formulation of the parameter estimation. Using the observation data determined by the experimental design, the calibrated unknown parameters become close to the actual value. Two case studies are provided to demonstrate the efficacy of the proposed method.

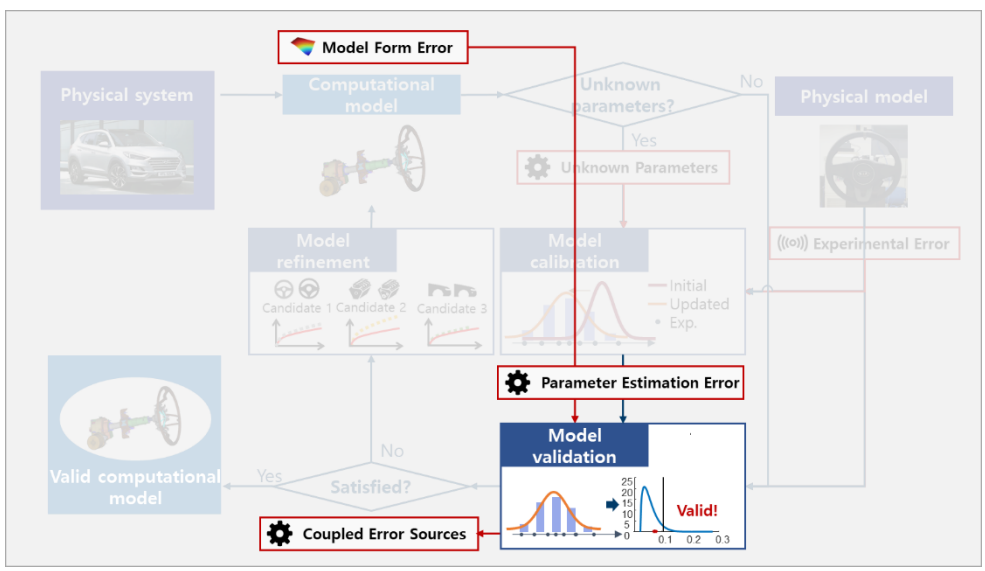


Figure 1-5 Model validation with coupled error sources

Research Thrust 2: Proportionate Bias Calibration with Bound Information to Consider Unrecognized Model Form Errors

The model calibration method with experimental design encounters a limitation in that the calibrated unknown parameters from the experimental design still remain parameter estimation errors (Figure 1-6). Even though the model calibration with experimental design approach can alleviate the estimation errors in the model calibration, it does not guarantee the optimal value of the unknown parameters. If the model form errors are dominant in the model calibration, experimental design cannot significantly reduce the amount of parameter estimation errors. These wrong estimated parameters can be irrational for practical engineering fields, which have expert-based information about these unknown model parameters. Thus, Research Thrust 2 focuses on a systematic framework for model calibration to consider the amount of unrecognized model form errors using the expert-based bound information. The reasonable bound information of unknown model parameters can be a bound constraint in OBMC. The proportionate bias represents the amount of model form errors in a relationship between the observation data and the predicted responses. The assumption of proportionate bias is that the biased error is proportional to the amount of predicted responses. The formulation of modeling bias is categorized into sensitivity-based bias and proportionate bias for multiple responses with severely different dimensions. Using the bias form, the amount of biased errors that arise due to unknown model form errors is calibrated with unknown model parameters. The bound information of unknown parameters can be a guide for OBMC to estimate a reasonable value of the unknown parameters and proportionate bias.

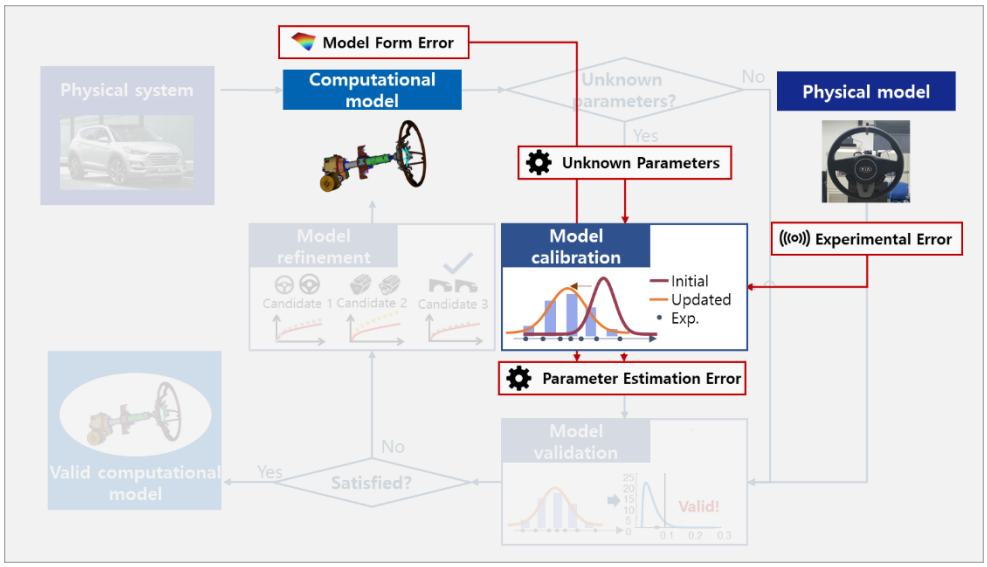


Figure 1-6 Parameter estimation errors that arise due to the main error sources in the computational and physical models

Research Thrust 3: Comparison of Statistical Validation Metrics to Reduce Type II Errors in Model Validation

Model validation is the process of determining the degree to which a computational model is an accurate representation of the actual 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 et al. 2011) (Ling and Mahadevan 2013) (Sankararaman and Mahadevanb 2015). In model validation, hypothesis testing is used to determine the validity or invalidity of a prediction based on the value of a statistical validation metric that quantifies the difference between the predicted and observed results. Errors in hypothesis testing decisions are troublesome when evaluating the accuracy of a computational model,

since an invalid model can be used in practical engineering design activities, and incorrect results in these settings may lead to safety issues. An appropriate selection of statistical validation metrics that are sensitive to the discrepancy of two distributions is required to reduce the decision errors. Thus, the objective of Research Thrust 3 is to provide a guideline to select reasonable statistical validation metrics. The decision errors of six statistical validation metrics, including area metric, Bayes factor, Likelihood, Kullback-Leibler Divergence, Probability of Separation, and Probability Residual, are compared with a numerical example with statistical and unknown model parameters. These comparison results propose the mean-supportive and variance-supportive metrics, according to whether the statistical validation metric is sensitive to the discrepancy of the mean or variance.

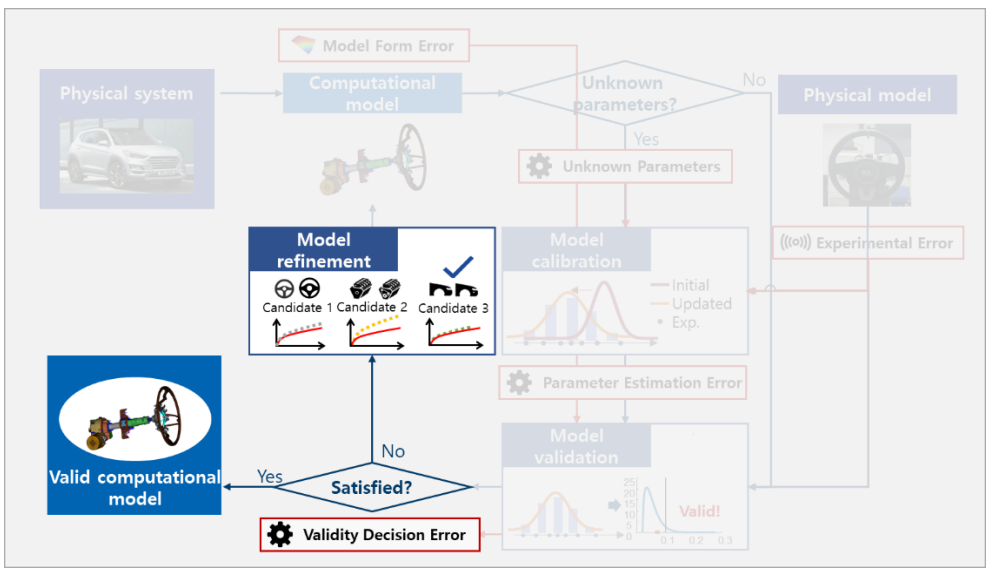


Figure 1-7 Validity decision errors in model validation

1.3 Dissertation Layout

This doctoral dissertation is organized as follows. Chapter 2 briefly explains overall process of optimization-based model improvement and its three sub-steps (e.g., model calibration, model validation, and model refinement). Chapter 3 presents an OBMI process that is integrated with experimental design, which reduces parameter estimation errors in model calibration. Chapter 4 proposes a new framework for model calibration with proportionate bias to consider model form uncertainties with a prior bound constraint of unknown model parameters. Chapter 5 suggests suitable statistical validation metrics from the perspective of mean and variance discrepancy to reduce type II errors. Chapter 6 summarizes the doctoral dissertation and its contributions and suggests future research directions. Appendix A provides an analytical representation of the statistical validation metrics considered in Chapter 5, when the interested responses are the normal and lognormal distribution. The derivation in Appendix A is used to explain the mean-supportive and variance-supportive characteristics of the validation metrics outlined in Chapter 5.

Chapter 2

Literature Review: Optimization-based and Bayesian-based Model Improvement

This chapter reviews two sorts of model improvement methods to help readers' understanding. Chapter 2.1 explains optimization-based model improvement (OBMI), the main interest of this doctoral dissertation. The subchapter includes the description of three sub-steps: model calibration, model validation, and model refinement. Chapter 2.2 introduces Bayesian-based model improvement process, another popular model improvement in a statistical manner. The comparison of these two methods are summarized in Chapter 2.3.

2.1 Optimization-based Model Improvement (OBMI)

Optimization-based model improvement consists of three processes: model calibration, model validation, and model refinement. Figure 1-1 illustrates overall process of OBMI. In If the modelers recognize the existence of unknown parameters, model calibration is performed first. The observation data from the experiments

plays a key role in calibration, serving as a representative of the true value. However, estimates of the unknown parameters can be biased from the true values due to error sources in both observation and prediction. Using the calibrated parameters, the model validation step is used to determine the validity of the computational model. Model validation uses a validation metric to measure the discrepancy between the observed and predicted results (Liu et al. 2011). When a computational model is invalid, modeling errors in the computational model are regarded as the main source of invalidity. Because the modelers typically believe that the computational model has been implemented with a sufficient amount of modeling knowledge, most modeling errors in a computational model remain unrecognized before the model improvement process. If the model validation step evaluates the predicted response from the computational model as invalid, the model improvement process performs model refinement to discover unrecognized modeling errors. Among all uncertainty candidates, model refinement distinguishes the most critical error sources through a three-step process: model invalidity analysis, an invalidity reasoning tree, and invalidity sensitivity analysis (Oh et al. 2016). The overall model improvement process continues until the model validation step satisfies the validity criterion.

2.1.1 Model Calibration

The model calibration is to find optimal values of unknown input parameters in a computational model (Trucano et al. 2006) (Arendt et al. 2010) (Youn et al. 2011) (Arendt et al. 2012) (Li et al. 2016) (Lee et al. 2018) (Jiang et al. 2020). In this approach, an optimization problem is formulated to estimate the unknown model parameters that minimize or maximize the discrepancy of the experimental and

computational responses (Frank and Shubin 1992) (Hills and Trucano 2002) (Gholizadeh and infrastructures 2013) (Lee et al. 2019). Figure 2-1 illustrates the overall process of optimization-based model calibration.

In Figure 2-1, the first step is to define the unknown model parameters. Generally, the existence of unknown model parameters are determined according to a modeler's opinion. However, the usual practical problems in model calibration encounter the ambiguity of how the prior information about model parameters is reliable. Thus, a variable selection method, such as sensitivity-based variable screening, can be adopted (Hamby 1994) (Frey and Patil 2002) (Campolongo et al. 2007) (Iooss and Lemaître 2015). With the unknown and known model parameters, the predicted models calculate the responses (outputs) to quantify the discrepancy from the observation data. The metric for quantification of discrepancy is called 'calibration metric'. The value of the calibration metric can be an objective function in the optimization problem.

$$\min_{\boldsymbol{\theta}} \text{ or } \max_{\boldsymbol{\theta}} f_{cal.}(\boldsymbol{\eta}(\boldsymbol{\theta}, \boldsymbol{x}; \mathbf{D}), \boldsymbol{\zeta}(\boldsymbol{x}; \mathbf{D})) \tag{2.1}$$

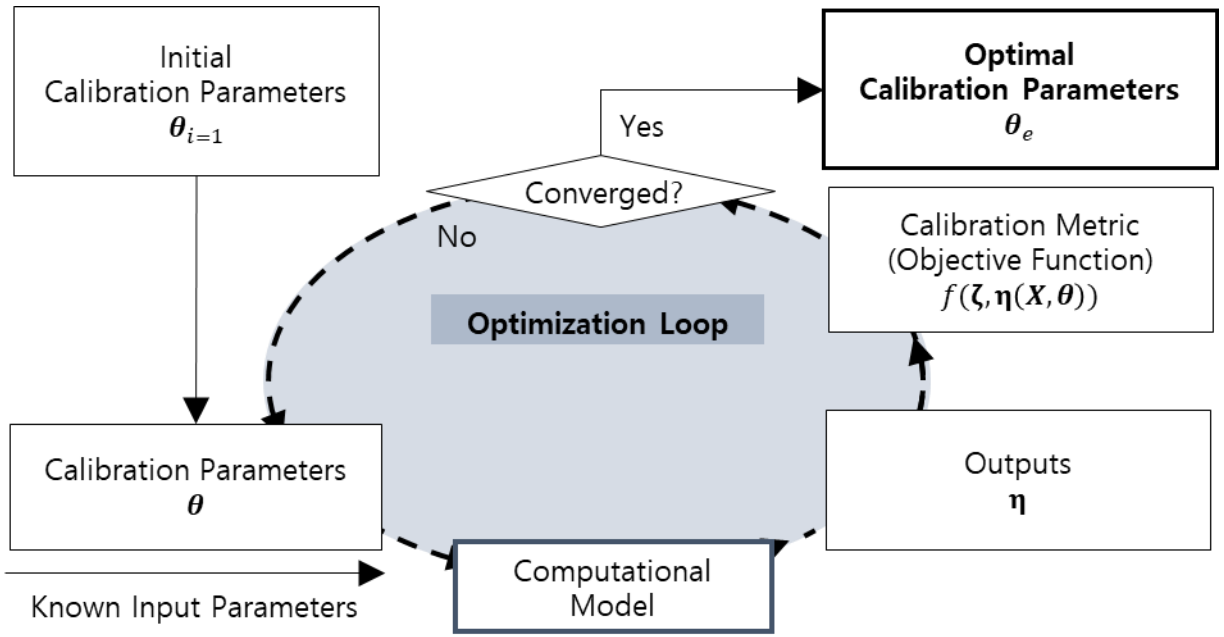


Figure 2-1 Optimization-based Model Calibration

In (2.1), f is a calibration metric, $\boldsymbol{\eta}$ is a predicted responses, and $\boldsymbol{\zeta}$ is observation data. In $\boldsymbol{\eta}$ and $\boldsymbol{\zeta}$, the terminologies $\boldsymbol{\theta}, \boldsymbol{x}$, and $\mathbf{D}$ respectively denote for unknown parameter, known parameter, and observation site. The OBMC can be performed deterministic or statistically, by adopting a calibration metric which can deal with deterministic or statistic values. The statistical calibration metrics utilizes a statistical moment or a probability density function of observation and prediction. This optimization loop continues until the calibration metric value satisfies the criteria pre-determined by the modelers.

It is worth noting that the optimization-based approach is a deterministic process since the optimal solution derived from the optimization-based approach is a deterministic value. However, the model improvement method in this paper is defined as a statistical approach because the method deals with uncertainties of responses over the whole process. When computational models include model form uncertainties, which arise mainly due to incorrect assumptions or excessive simplification, model calibration fails to find reasonable estimates of the statistical moments of the unknown input variables. The model refinement step can revise the leading cause of the model form uncertainties. To avoid a failure of model calibration, exact computational modeling and simplification of assumptions based on reasonable physics are required.

2.1.2 Model Validation

Model validation is a process to verify the validity of computational responses (Oberkampf and Roy 2010). It consists of examining the validation metric and decision making. The validation metric quantifies the coincidence or the discrepancy

between the experimental and computational responses. The validation metric is technically different from the calibration metric in that it should give information of model validity even in constrained cases, such as where there is a limited number of data or where the experimental data is given in different environments. To consider the statistical uncertainty in limited observation, most of model validation adopts a statistical validation metric. Statistical validation metrics include the area metric and Bayes factors (Oberkampf and Trucano 2002) (Rebba et al. 2006) (Liu et al. 2011). Using the validation metric value, the model validation is used to determine the validity of the calibrated model. In decision making, hypothesis testing is generally used.

Model validation includes the comparison of a model prediction with experimental data to evaluate a computational model's prediction accuracy. Figure 2-2 shows the flow chart of the overall model validation process (Oberkampf and Barone 2006) (Kat and Els 2012). The first step of model validation is to prepare the prediction and observation results using computational models and physical experiments. Based on this prediction and observation data, statistical validation metrics quantify the discrepancy between the computational prediction and experimental observation. Depending on The validation metric result composes a deterministic quantity that is then assessed in model validation. This study distinguishes statistical validation metrics as either larger-the better (L type) or smaller-the better (S type). L-type metrics increase when the prediction and observation become more similar. S-type metrics decrease as the prediction and observation become more similar. Chapter 5.1 provides details about the characteristics of each statistical validation metric.

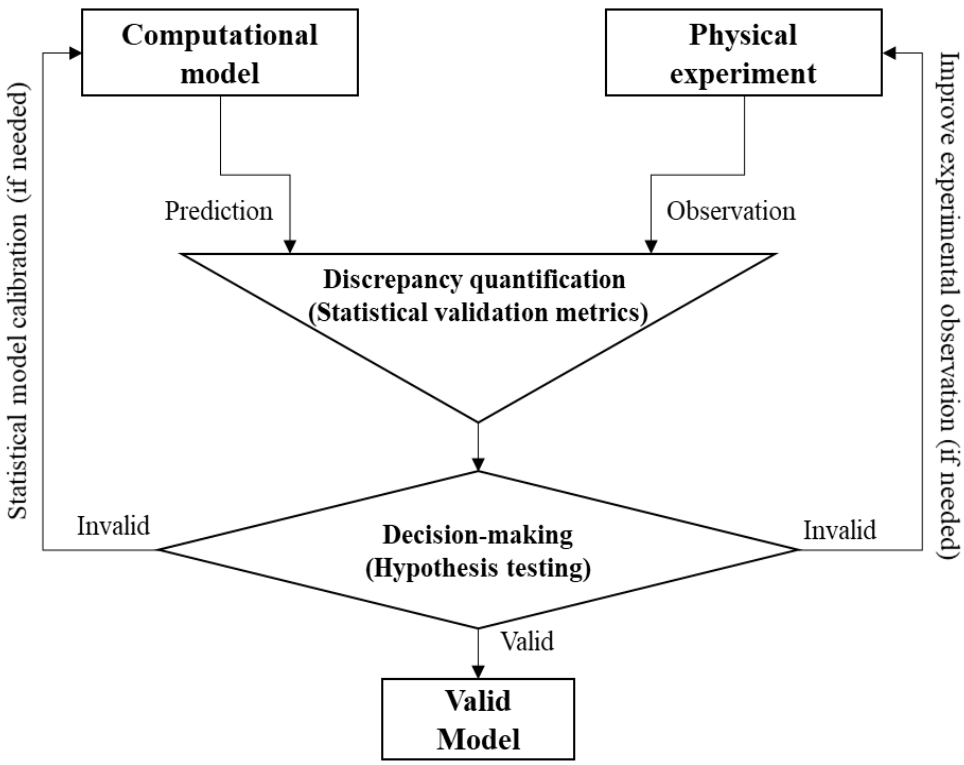
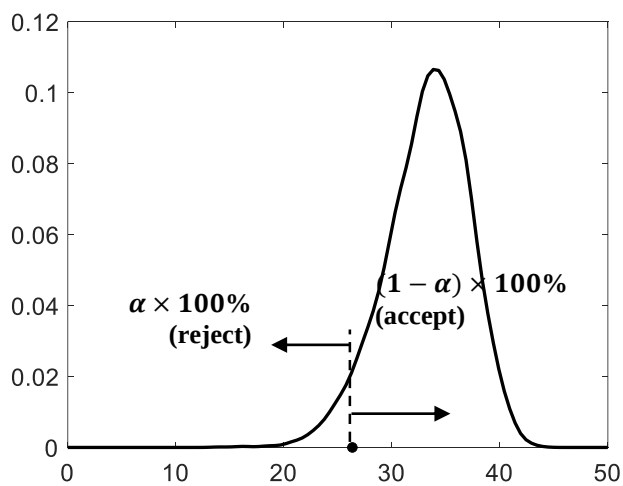
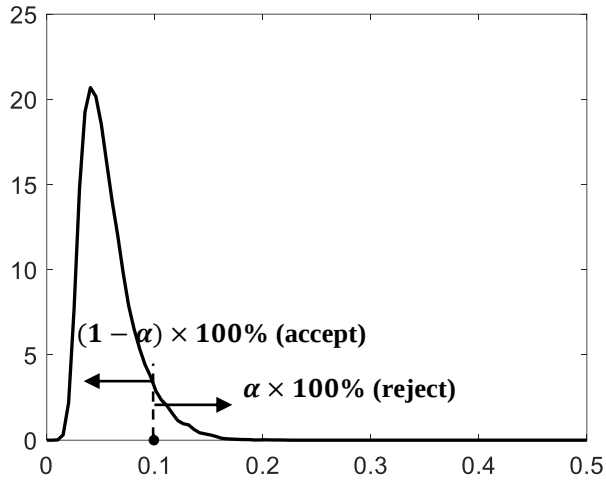


Figure 2-2 The process of statistical model validation (SMV).

To evaluate the prediction, hypothesis testing examines the probability of the null hypothesis and an alternative hypothesis, using the statistical validation metrics. The key idea is to determine the criteria by building a distribution of the null hypothesis in which the prediction model is valid. To do this, first, numerous observation datasets are randomly sampled under the distribution fitted by the prediction. The validation metric calculated by the observation datasets then forms the distribution of the null hypothesis. The confidence level (denoted by α) distinguishes the ranges in which the null hypothesis should be accepted or rejected for a validation metric value. Figure 2-3 explains the location of the confidence level in the validation metric distribution. Figure 2-3(a) shows the distribution of L-type metrics, such as likelihood and Bayes factor. The distribution in Figure 2-3(b) is given for S-type metrics, including the area metric, Kullback-Leibler divergence, probability of separation, and probability residual.



(a)



(b)

Figure 2-3 The confidence level in a validation metric's distribution (a) L-type (log-likelihood), (b) S-type (area metric with U-pooling).

2.1.3 Model Refinement

Model refinement is an essential step in that this process directly removes the underlying cause of unrecognized model form errors. There has been little achievement in academic research towards considering the unrecognized model form errors in a systematic framework. Xiong et al. stated in their paper that personal experience in modeling gives an intuition for model refinement: however, this is not an applicable process for practical settings. Therefore, model refinement, introduced in this paper, aims to explore the most effective root causes of invalid modeling via a systematic approach (Oh et al. 2016) .

The model refinement selects the most invalid sources based on experts'

opinion and numerically quantified criteria. The process involves three steps: 1) model invalidity analysis, 2) an invalidity reasoning tree, and 3) invalidity sensitivity analysis. Model invalidity analysis is a brainstorming step, which gathers all possible invalid sources. This step allows as many invalid sources as possible. Figure 2-4 shows an example of the affinity diagram for model invalidity analysis. The second step is to develop an invalidity reasoning tree that selects only potential invalid sources from among all sources gathered in Step 1. Figure 2-5 shows an example of invalidity reasoning tree. For the selection of invalid sources, related experts should identify the proper reasons for invalidity, from the conceptual, mathematical, and computational perspective. Invalidity sensitivity analysis quantitatively evaluates the importance of invalid sources of computational modeling. A decision matrix is one useful tool for comparing all candidates of the invalid sources of modeling (Dieter 1991). Oh et al. suggested a weighted decision matrix that considers the importance of each criterion by multiplying the weight values (Forman and Gass 2001) (Oh et al. 2016). Figure 2-6 is an example of the objective tree which gives weight values for each criteria. Figure 2-7 shows an example of weighted decision matrix. In this step, engineers in the field can define criteria and weights for quantification appropriate for their situation.

Constraint	- Constraint location of airbag is different from a real specimen - Rigid constraint of an airbag does not reflect reality
Physics	- Shell(Plate) assumption of wheel cover is not proper
Geometry	- Modeling Simplification of an airbag as a lumped mass

Figure 2-4 Example of affinity diagram for model invalidity analysis

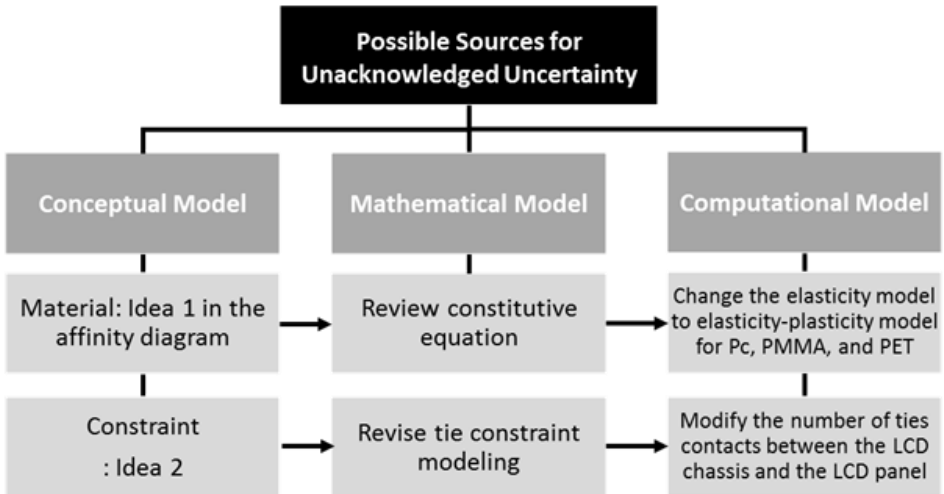


Figure 2-5 Example of invalidity reasoning tree

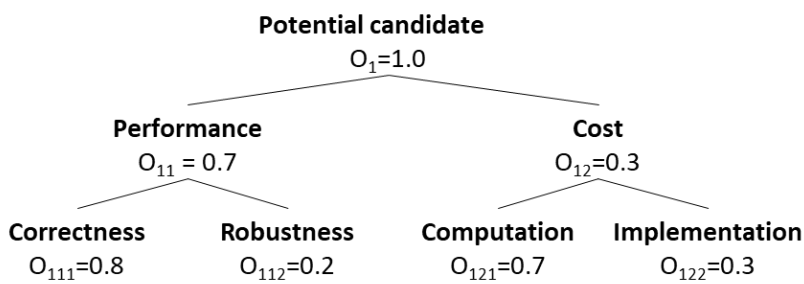


Figure 2-6 Example of objective tree with weight values

Criteria	Weight	Candidate 1		Candidate 2		Candidate 3	
		Score	Rating	Score	Rating	Score	Rating
Correctness	0.48	5	2.40	1	0.48	4	1.92
Robustness	0.12	3	0.36	3	0.36	3	0.36
Computation cost	0.28	3	0.84	4	1.12	3	0.84
Implementation cost	0.12	2	0.24	4	0.48	4	0.48
Sum of rating			3.84		2.44		3.60

Figure 2-7 Example of invalidity sensitivity analysis

Numerous field engineers use simplified computational models to increase the speed of calculation or the efficiency of the modeling. The majority of modeling errors arise from the simplification of computational models, rather than a lack of modeling knowledge. Therefore, invalid modeling should be selectively improved by considering the situation and the standards required for each industrial field to ensure that an unconditionally exact and complicated model is not implemented. In addition to the model refinement method, other statistical approaches to deal with the invalidity of a computational model have been developed (Kennedy and O'Hagan 2001) (Xiong et al. 2009) (Qiu et al. 2018). Bias correction with Bayesian-based model improvement framework can quantify the number of errors that are due to the invalidity of a computational model (Kennedy and O'Hagan 2001) (Arendt et al. 2012) (Xi et al. 2013). To precisely quantify the biased errors, this method requires numerous experimental data samples from a diverse domain of design variables. The following chapter 2.2 provides more details of bias correction method in Bayesian-based model improvement.

2.2 Bayesian-based Model Improvement with Bias Correction

Bayesian-based model calibration is another popular approach for model improvement which adopts Bayesian inference to estimate unknown model parameters (Kennedy and O'Hagan 2001) (Beck and Au 2002) (McFarland et al. 2007) (Higdon et al. 2008) (Behmanesh et al. 2015) (Li et al. 2016) (Plumlee 2017) (Baig 2020). Figure 2-8 illustrates the overall flow chart of Bayesian-based model improvement. This process applies observation data and the prior distribution of

unknown model parameters to Bayesian inference, stated in (2.2).

$$p(\boldsymbol{\theta}|\boldsymbol{\eta}) \propto p(\boldsymbol{\eta}|\boldsymbol{\theta})p(\boldsymbol{\theta}) \tag{2.2}$$

The term $p(\boldsymbol{\theta}|\boldsymbol{\eta})$ is a posterior distribution of unknown model parameters θ. The term $p(\boldsymbol{\eta}|\boldsymbol{\theta})$ is likelihood function for unknown model parameters. The last term $p(\boldsymbol{\theta})$ is a prior distribution of unknown parameters.

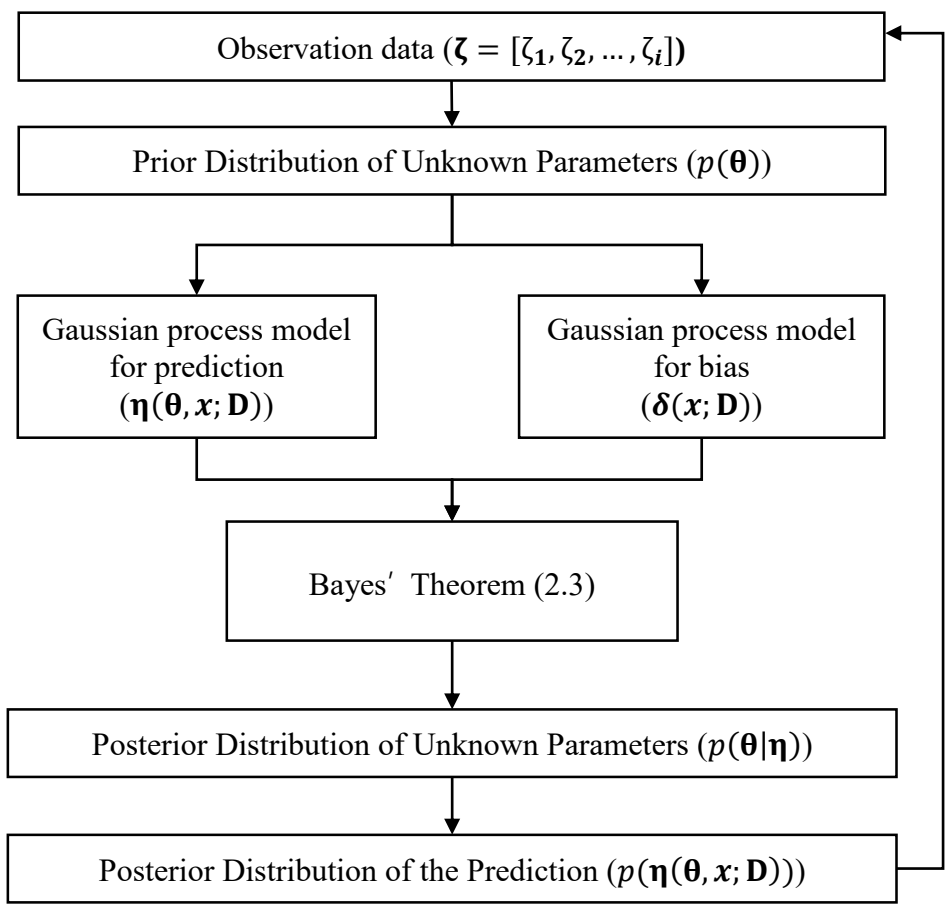


Figure 2-8 The flow chart of Bayesian-based model improvement

The prediction model $\boldsymbol{\eta}$ in (2.2) can be defined under the effect of model form errors and measurement errors, as shown in (2.3).

$$\begin{aligned} \boldsymbol{\eta}(\boldsymbol{\theta}, \boldsymbol{x}; \mathbf{D}) + \boldsymbol{\delta}(\boldsymbol{x}; \mathbf{D}) &= \boldsymbol{\zeta}(\boldsymbol{x}; \mathbf{D}) + \boldsymbol{\varepsilon}(\mathbf{D}) \\ \boldsymbol{\eta}(\boldsymbol{\theta}, \boldsymbol{x}; \mathbf{D}) &= \boldsymbol{\zeta}(\boldsymbol{x}; \mathbf{D}) - \boldsymbol{\delta}(\boldsymbol{x}; \mathbf{D}) + \boldsymbol{\varepsilon}(\mathbf{D}) \end{aligned} \tag{2.3}$$

where $\boldsymbol{\delta}$ is the unrecognized source of modeling errors and $\boldsymbol{\varepsilon}$ is measurement error. One thing to be clarified is that the exact amount of error sources $\boldsymbol{\delta}$ and ε are unknown. Bayesian-based method employs bias correction to quantify the amount of modeling errors using observations measured across various sites in the design parameters (Rebba et al. 2006) (Higdon et al. 2008) (Arendt et al. 2010) (Li et al. 2016). The bias correction method first quantifies the bias term as the discrepancy between prediction and observation with the prior information of unknown parameters. And then, the quantified bias is adopted to (2.3) for Bayesian inference to estimate the posterior distribution of unknown parameters and prediction. Gaussian process modeling can deal with the uncertainty in bias and prediction response. Generally, measurement error is assumed to be a normal distribution, with zero-mean and covariance estimated by the observation data (Ferson et al. 2008).

Even though the bias correction method is an efficient way to quantify the model form errors and unknown parameters together, neither values are similar to the actual value (Oliver et al. 2015). The bias term determined by this framework depends on a prior distribution of unknown model parameters, which affects the result of Bayesian calibration (Jiang et al. 2020). In other words, the result of Bayesian calibration differs from how accurate the prior information is. The aims of

quantifying bias term to minimize the difference between the observed and predicted results. Based on the observation data obtained from design parameters from various sites, the bias and estimated unknown parameters can be updated to represent the prediction values. Overall, this technique has no need to consider the accuracy of parameter estimation errors or the effect of invalidity that arises from modeling errors; however, this technique requires numerous observation data.

2.3 Summary and Discussion

Chapter 2 aims to introduce two sorts of popular model improvement process: 1) optimization-based approach, and 2) Bayesian-based approach. It differs in what kinds of the methods the model improvement framework employs for model calibration. By the explanation described in Chapter 2.1 and 2.2, Table 2-1 compares the characteristics of these two model improvement process.

OBMI is a deterministic method that gives a deterministic value for estimated unknown parameters. The method can deal with the uncertainties of unknown parameters as a statistical moment (e.g., mean and standard deviation). Furthermore, OBMI allows getting an optimal value of unknown parameters without any information about model parameters. OBMI has initially aimed for a computational model to assist an engineered product design, which has difficulty measuring observation data for various design sites. Thus, three steps in OBMI aim to enhance a computational model's accuracy with a limited or insufficient amount of observation data. For example, the OBMI devices model refinement step which handles unrecognized model form errors in a qualitative approach.

Table 2-1 Comparison of optimization-based model improvement and Bayesian-based model improvement

		Optimization-based Model Improvement	Bayesian-based Model Improvement
Method		Deterministic	Statistic
Output		A representative value (Mean, standard deviation)	Probability distribution
Error Source Consideration	Experimental error (ε)	Assumes the error is too small to ignore	Assumes normal distribution with zero-mean $N\sim(0,\sigma^2)$
	Model form error (δ)	Explore the error sources by model refinement method	Gaussian process modeling with bias correction method
Required prior information		-	distribution of unknown model parameters
Applicable Observation Data		Data measured at limited design sites	Datasets measured in a diverse design sites

Compared with OBMI, Bayesian-based model improvement is a statistical approach that gives a result of model calibration as a probability distribution. The method is an efficient framework to apply the prior information of the parameters. However, the result of estimated unknown parameters highly depends on the accuracy of the prior information. With bias correction method, Bayesian-based model improvement can quantify the Gaussian process model for model form errors about design parameters. Thus, the Bayesian-based approach requires is applicable for an engineering system which requires system health diagnostics and management.

By the above comparison, the notable difference between OBMI and Bayesian-based model improvement is the amount of information used for model improvement. OBMI intends to be capable of model improvement process when the applicable information is limited. In contrast, Bayesian-based model improvement aims to appropriate use of assorted information such as a prior distribution of unknown parameters or numerous measurable observation datasets. This doctoral dissertation aims to develop a capable model improvement process in most practical situations where datasets are limited in a system design step. The research topics introduced in Chapter 3 through Chapter 5 were devised under the assumption that only a partial dataset measured at the specific design conditions are allowable.

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

1) **Hyejeong Son**, Guesuk Lee, Kyeonghwan Kang, Young-jin Kang, Byeng D. Youn, Ikjin Lee, and Yoojeong Noh, "Industrial Issues and Solutions to Statistical Model Improvement: A Case Study of an Automobile Steering Column," *Structural and Multidisciplinary Optimization*, Vol. 61, pp. 1739-1756, 2020.

Chapter 3

Experimental Design for Identifying Error Sources Between Parameter Estimation Errors and Model Form Errors

Numerous error sources in observation and prediction affect the results of model improvement. Observation involves measurement errors, which arise due to environmental noise, lack of precision in the sensing, or an incorrect experimental setup (Kim et al. 2019). In a computational model, insufficient knowledge, or excessive simplification, induces modeling errors that degrade the prediction accuracy (Thakur et al. 2009). These error sources can degrade the accuracy of the parameter estimation step in model calibration (Oberkampf and Roy 2010) (Lee et al. 2019) (Lee 2019). These parameter estimation errors can be another source of invalidity, in addition to the modeling errors in a computational model. Thus, the model improvement process requires consideration of the effects of error sources from both observation and prediction.

To address modeling errors of a computational model in OBMI, model refinement is used to discover unrecognized sources of errors in the model. Model

refinement is conducted after the model validation step only when the predicted responses of a computational model show invalidity in the model validation step (Oh et al. 2016) (Son et al. 2020). This method is able to apply model improvement with limited observation data in a specific design domain, ignoring measurement errors. However, the parameters estimated from this type of model calibration are difficult to apply under various design domain conditions, due to parameter estimation errors. If the model validation step determines that the computational model is invalid, either modeling errors or parameter estimation errors can be the main cause of invalidity. Since the model refinement step cannot deal with both parameter estimation errors and modeling errors, this method is limited in its ability to improve the accuracy of the overall prediction.

The research outlined herein develops an optimization-based model improvement that distinguishes the sources of prediction invalidity, discriminating between errors that arise from the calibrated parameters and those from the modeling errors in a computational model. An experimental design is integrated into the optimization-based model improvement to reduce the estimation errors that arise from unknown parameters in the model calibration step. Using this experimental design, the errors that remain in a computational model after model calibration are then mainly due to unrecognized sources of modeling error. Thus, the refinement step can focus on reducing these unrecognized sources of modeling errors.

The remaining chapters are organized as follows. Chapter 3.1explains the model improvement process, considering the various types of uncertainty. Chapter 3.2 provides an overview of the proposed model improvement process, including information-based experimental design. Chapter 3.3 discusses the numerical and

practical case studies that were examined in this work to show the validity of the proposed method. Finally, Chapter 0 concludes with overall remarks and a discussion of future work.

3.1 Coupled Error Sources in Model Calibration

The flow chart in Figure 3-1 shows optimization-based model improvement, which was developed to deal with error sources in observation and prediction (Youn et al. 2011). From a physical system in which an interested response can be observed, a modeler constructs an initial computational model based on the available knowledge. Observation and prediction respectively include measurement errors and model form errors. As mentioned in Chapter 2, OBMI process generally assumes that the observational result are free from any errors and negligible. Modeling errors can result from any of the diverse types of knowledge required for modeling, such as input parameters for which exact values are unknown, or invalid assumptions in a model form. The input parameters denote the physical quantities that determine the characteristics of a system, such as geometric figures, material properties, or load and boundary conditions in a computational model. Modeling assumptions involve geometric simplification, linearization, or elasticity assumptions that are employed for computational convenience.

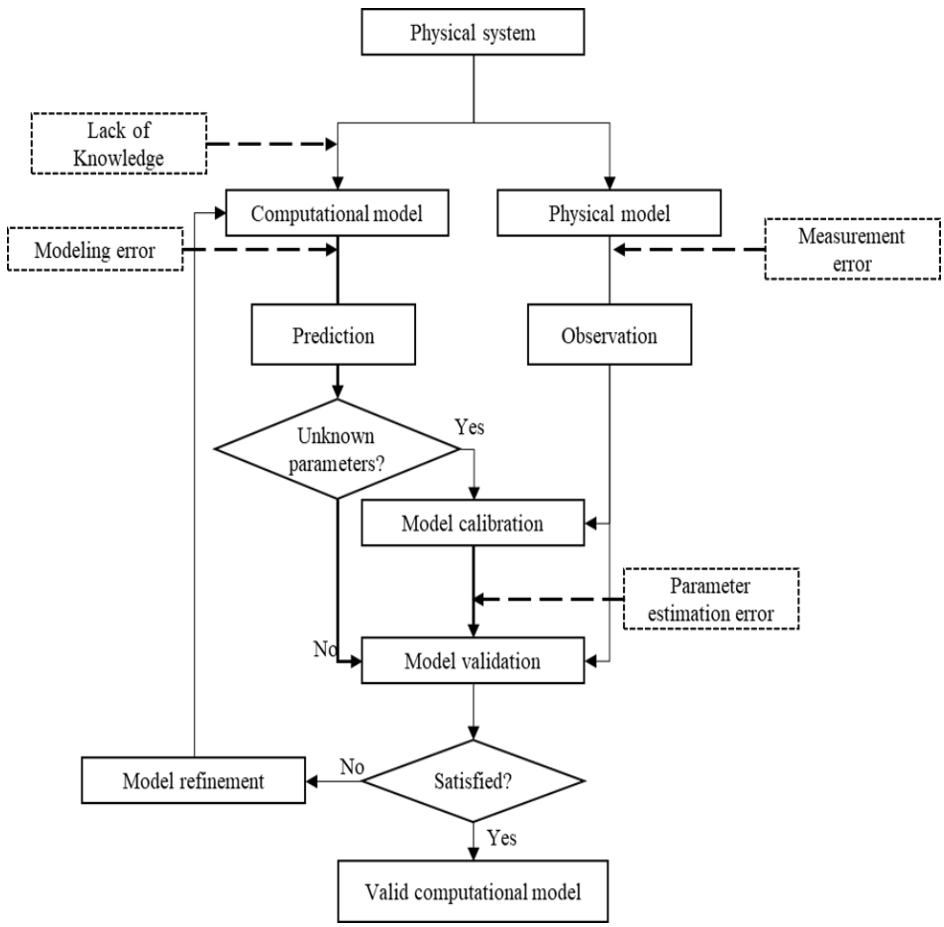
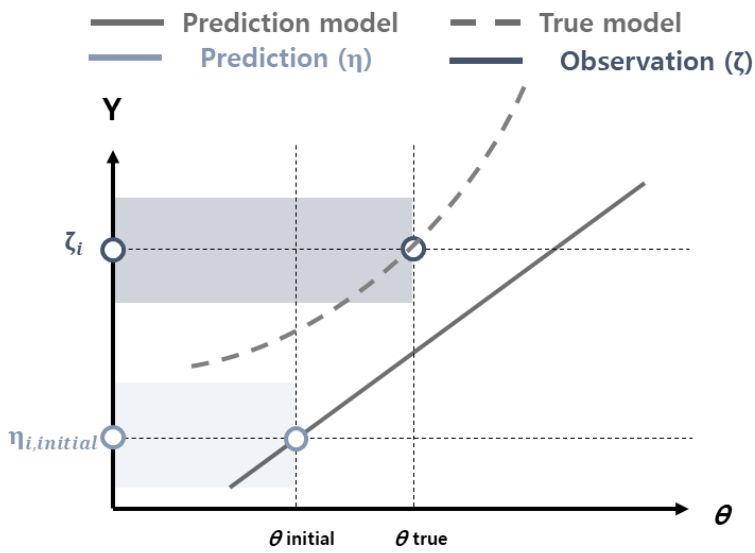
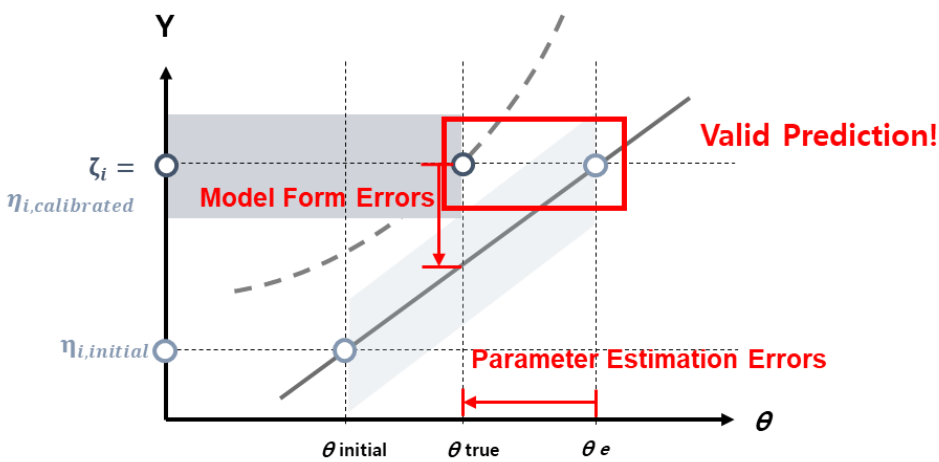


Figure 3-1 Model improvement process that considers major error sources

The existence of model form errors is a nuisance for model calibration in that it affects the accuracy of estimated unknown parameters and has difficulty quantifying the exact amount of errors. Figure 3-2 illustrates the model calibration process of a prediction model with model form errors. Figure 3-2 (a) shows the prediction and observation before the model calibration. The true model is assumed as a linear function in the prediction model. When $\boldsymbol{\theta} = \theta_{\text{initial}}$, initial prediction ($\eta_{i,initial}$) is different from observation (ζ_i). The model calibration tries to optimize θ which makes prediction and observation equal ($\eta_i = \zeta_i$). After the model calibration, the figure changes to Figure 3-2 (b). Even though the prediction and observation responses became similar, θ is optimized to θ_e, biased from θ_{true} due to model form errors. Thus, the model validation step cannot evaluate the invalidity of the prediction model since the errors of unknown parameters supplement the errors in prediction.



(a)



(b)

Figure 3-2 Model calibration with model form errors: (a) Before calibration, (b) After calibration

3.2 Optimization-based Model Improvement with Experimental Design

Chapter 3.1 described the primary limitation of optimization-based model improvement: the absence of an ability to identify the sources of uncertainty. Thus, this study proposes a method that helps to identify the uncertainty sources by reducing the calibration errors. For this purpose, an experimental design method is adopted. This chapter explains the effects that error sources in observation and prediction have on model calibration. Then, an experimental design process is introduced to reduce the errors in the estimated unknown parameters.

3.2.1 Derivation of Parameter Estimation Errors in Model Calibration

There are basically two error sources that degrade the accuracy of a prediction: parameter estimation errors from the model calibration step and unrecognized modeling errors in the computational model. Since these two error sources simultaneously affect the computational model, it is difficult to distinguish which is the main error source when the target model turns out to be invalid. In this study, we propose a method that helps to identify the uncertainty sources by minimizing the effect of the parameter estimation errors. When parameter estimation errors are minimal, the dominant invalidity sources that remain after calibration are the unrecognized modeling errors; thus, the model refinement process can focus on improving the modeling error. To discuss the details of the proposed method, this study first explains how measurement errors and modeling errors affect the estimates of unknown parameters in the optimization-based model improvement process.

The general equation of a predicted and observed response, which considers the error sources in observation and prediction, is expressed as (3.1) (Kennedy and O'Hagan 2001) (Xiong et al. 2009) (Arendt et al. 2012) (Plumlee 2017).

$$\boldsymbol{\eta}(\boldsymbol{\theta}, \boldsymbol{x}; \mathbf{D}) + \boldsymbol{\delta}(\boldsymbol{\theta}, \boldsymbol{x}; \mathbf{D}) = \boldsymbol{\zeta}(\boldsymbol{x}; \mathbf{D}) + \boldsymbol{\varepsilon}(\mathbf{D}) \tag{3.1}$$

where $\boldsymbol{\eta}$ is the predicted response of a computational model and $\boldsymbol{\zeta}$ is the target or true response of a physical model. $\boldsymbol{\delta}$ is the unrecognized source of modeling errors and $\boldsymbol{\varepsilon}$ is measurement error. $\boldsymbol{x}$ and $\boldsymbol{\theta}$ stand for design parameters and unknown parameters, respectively. The design parameter $\boldsymbol{x}$ is controllable in a design process and affects the predicted responses. In the optimization-based model improvement process, however, $\boldsymbol{x}$ can be treated as a constant parameter because the process is generally employed in a specific design domain (Youn et al. 2011) (Jung et al. 2015). $\mathbf{D}$ represents the experimental point where the responses are measured. In (3.1), the measurement errors $\boldsymbol{\varepsilon}$ are independent from other input parameters and only related to the measurement location. The unrecognized modeling error $\boldsymbol{\delta}$ depends on the design parameter $\boldsymbol{x}$ and the measurement location $\mathbf{D}$. In the model improvement process, the first step is model calibration to estimate $\boldsymbol{\theta}$, using the predicted response and observation. Model calibration is an inverse problem, which solves the equation for $\boldsymbol{\theta}$. To derive the estimates of $\boldsymbol{\theta}$, linear approximation of the predicted response $\boldsymbol{\eta}$ is performed, as follows.

$$\boldsymbol{\eta}(\boldsymbol{\theta}, \boldsymbol{x}; \mathbf{D}) \approx \boldsymbol{\eta}(\boldsymbol{\theta}_{\boldsymbol{a}}, \boldsymbol{x}; \mathbf{D}) + (\boldsymbol{\theta} - \boldsymbol{\theta}_{\boldsymbol{a}})\nabla_{\boldsymbol{\theta}}\boldsymbol{\eta} \tag{3.2}$$

$\nabla_{\boldsymbol{\theta}}\boldsymbol{\eta}$ denotes a (m×n) gradient matrix of $\boldsymbol{\eta}$ which has m number of responses with regard to the n number of unknown parameters. $\boldsymbol{\theta}_{\mathbf{a}}$ is an arbitrary value of the

unknown parameters within a prior bound of unknown parameters. By substituting (3.2) into (3.1), the equation of prediction and observation is as follows.

$$\boldsymbol{\eta}(\boldsymbol{\theta}_a, \boldsymbol{x}; \mathbf{D}) + (\boldsymbol{\theta} - \boldsymbol{\theta}_a)\boldsymbol{\nabla}_{\boldsymbol{\theta}}\boldsymbol{\eta} + \boldsymbol{\delta}(\boldsymbol{\theta}, \boldsymbol{x}; \mathbf{D}) = \boldsymbol{\zeta}(\boldsymbol{x}; \mathbf{D}) + \boldsymbol{\varepsilon}(\mathbf{D}) \tag{3.3}$$

Solving this equation about $\boldsymbol{\theta}$, the estimate $\boldsymbol{\theta}_e$ for unknown parameter $\boldsymbol{\theta}$ is derived in (3.4).

$$\begin{aligned}\boldsymbol{\theta}_e = \{(\boldsymbol{\nabla}_{\boldsymbol{\theta}}\boldsymbol{\eta})^{\mathrm{T}}(\boldsymbol{\nabla}_{\boldsymbol{\theta}}\boldsymbol{\eta})\}^{-1}(\boldsymbol{\nabla}_{\boldsymbol{\theta}}\boldsymbol{\eta})^{\mathrm{T}}\{\boldsymbol{\zeta}(\boldsymbol{x}; \mathbf{D}_c) + \boldsymbol{\varepsilon}(\mathbf{D}_c) \\ - \boldsymbol{\delta}(\boldsymbol{\theta}, \boldsymbol{x}; \mathbf{D}_c) - \boldsymbol{\eta}(\boldsymbol{\theta}_a, \boldsymbol{x}; \mathbf{D}_c)\} + \boldsymbol{\theta}_a)\end{aligned} \tag{3.4}$$

Dc refers to the measurement location for model calibration. The pseudo inverse is adopted to calculate the inverse of $\boldsymbol{\nabla}_{\boldsymbol{\theta}}\boldsymbol{\eta}$, which can be a non-square matrix. If the predicted responses are nonlinear, the optimization algorithm can be used to find the optimal value of $\boldsymbol{\theta}_e$ that minimizes the discrepancy between observed and predicted results, rather than solving an analytical $\boldsymbol{\theta}_e$. No matter how the inverse problem is solved, $\boldsymbol{\theta}_e$ is deeply related to $\boldsymbol{\varepsilon}$ and $\boldsymbol{\delta}$. Thus, $\boldsymbol{\theta}_e$ cannot guarantee that the estimate is the same as the true $\boldsymbol{\theta}$. In the model validation step, the estimate $\boldsymbol{\theta}_e$ ensures validity only for designated design site $\boldsymbol{x}$ and location parameter $\mathbf{D}_c$. In order to satisfy validity in situations other than a specific calibration domain, it is of importance to minimize the parameter prediction errors in model calibration by employing an experimental design.

3.2.2 Identification of Error Sources by Employing Experimental Design

An accurate understanding of error sources in observation and prediction requires

enormous time and effort. Quantification of measurement errors require repeated experiments with precise testing equipment. Discovering all modeling errors is time-consuming because modelers must examine the majority of the diverse modeling conditions. Therefore, it is nearly impossible to directly identify and separate all sources of errors. As an efficient tool for identifying the error sources – specifically, distinguishing between calibrated errors and unrecognized modeling errors – this study proposes integration of the model improvement process with experimental design to reduce estimation errors that arise from unknown parameters in model calibration. The experimental design step finds the optimal location for observation, which can reduce the parameter estimation errors that otherwise arise in model calibration. After the model calibration step, the invalidity sources in the computational model that remain can be assumed to primarily be unrecognized modeling errors. Then, the model improvement process can ensure that the model refinement step can be used modify the prediction accuracy when the model validation step determines that the predicted response is invalid.

Experimental design processes employ numerous alphabetic criteria for optimizing the measurement location, such as A-, C-, E-, T- or D-optimality (Park and Himmelblau 1982) (Ucinski 2004) (Pukelsheim 2006) (Atkinson et al. 2007) (Tricaud et al. 2008) (Song et al. 2009) (Maes et al. 2015). The alphabetic criteria are used to minimize or maximize a matrix form using an invariant representative of a matrix (Bandara et al. 2009) (Bock et al. 2013) (Chisari et al. 2017). In this study, experimental design with a D-optimality condition is adopted. This optimality criterion is the minimization of the following determinant of dispersion matrix $\left|\mathbf{X}^{\mathrm{T}}\mathbf{X}\right|^{-1}$, as shown in (3.5) (de Aguiar et al. 1995).

$$M = \det\left(\left|\mathbf{X}^{\mathrm{T}}\mathbf{X}\right|^{-1}\right) \tag{3.5}$$

X denotes the gradient of the computational model (John and Draper 1975). In a statistical manner, **X** can be substituted into the likelihood of the unknown parameters (Bock et al. 2013) (Papadimitriou and Papadimitriou 2015). This approach has an advantage in multivariate model calibration, since minimal D-optimality can confirm independence among the measured responses. The minimal D-optimality condition guarantees non-trivial solutions in model calibration.

This study discusses how the D-optimality experimental design method can reduce the estimation errors in the unknown parameters. It starts with (3.2), which represents a linear approximation of a computational model. Assuming that the observation and prediction have no error sources, the true value of the unknown parameter is derived as outlined in (3.6).

$$\boldsymbol{\theta} = \{(\boldsymbol{\nabla_{\theta}\eta})^{\mathrm{T}}(\boldsymbol{\nabla_{\theta}\eta})\}^{-1}(\boldsymbol{\nabla_{\theta}\eta})^{\mathrm{T}}\{\boldsymbol{\zeta}(\boldsymbol{x};\mathbf{D_c}) - \boldsymbol{\eta}(\boldsymbol{\theta_a},\boldsymbol{x};\mathbf{D_c})\} + \boldsymbol{\theta_a} \tag{3.6}$$

From (3.4) and (3.6), the estimation error of the unknown parameters is calculated as follows.

$$\boldsymbol{\theta} - \boldsymbol{\theta_e} = \{(\boldsymbol{\nabla_{\theta}\eta})^{\mathrm{T}}(\boldsymbol{\nabla_{\theta}\eta})\}^{-1}(\boldsymbol{\nabla_{\theta}\eta})^{\mathrm{T}}\{\boldsymbol{\varepsilon}(\mathbf{D_c}) - \boldsymbol{\delta}(\boldsymbol{\theta_a},\boldsymbol{x};\mathbf{D_c})\} \tag{3.7}$$

Equation (3.7) indicates that the estimation error of the unknown parameters is proportional to the pseudo-inverse of gradient matrix $\boldsymbol{\nabla_{\theta}\eta}$, $\boldsymbol{\varepsilon}$, and $\boldsymbol{\delta}$. Among them, D-optimality is equal to minimization of the determinant of $\{(\boldsymbol{\nabla_{\theta}\eta})^{\mathrm{T}}(\boldsymbol{\nabla_{\theta}\eta})\}^{-1}$ in the pseudo inverse form. **X** in (3.5) corresponds to $\boldsymbol{\nabla_{\theta}\eta}$ in (3.7). The method ignores

the effect of $\boldsymbol{\varepsilon}$ and $\boldsymbol{\delta}$, since the exact error sources $\boldsymbol{\varepsilon}$ and $\boldsymbol{\delta}$ are unknown to the user. The effect of $\boldsymbol{\varepsilon}$ and $\boldsymbol{\delta}$ is practically uncontrollable. The minimization of D-optimality can reduce the parameter estimation errors in model calibration, even though the reduced amount of parameter estimation error does not actually reach its minimum due to the uncontrollable $\boldsymbol{\varepsilon}$ and $\boldsymbol{\delta}$. In the optimization-based model improvement process, this experimental design prevents model calibration from adjusting the unknown parameters to compensate for degraded prediction that arises due to modeling errors. Using estimates of the unknown parameters provided by the experimental design, the remaining error sources evaluated by the model validation are constrained to the unrecognized modeling errors. Ultimately, using this approach, model refinement can then efficiently explore the remaining unrecognized modeling errors without the interference of other error sources.

3.3 Case Studies

Two case studies are provided to demonstrate the validity of the proposed method of model improvement with the D-optimality based experimental design method. Chapter 3.3.1 gives an analytical discussion of model improvement with D-optimality based experimental design to study a cantilever beam. Chapter 3.3.2 provides a practical example from an engineering field that examines an automotive wheel rim model. These two examples consider different types of unrecognized error sources to demonstrate the usefulness of the proposed method in different situations.

3.3.1 Analytical Case Study: Cantilever Beam Model

In this case study, a simple cantilever beam model is adopted to demonstrate the proposed method. The beam is subjected to a concentrated load at the tip and rectangular section. The model satisfies the Euler-Bernoulli beam theory. The setup is shown in Figure 3-3. Here, we assume that some parameters of the cantilever beam are estimated using the stress and displacement, whose equations are given as follows.

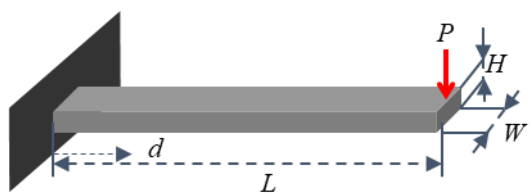


Figure 3-3 A cantilever beam

$$\boldsymbol{\eta_{true}} = \begin{bmatrix} \sigma_{true} \\ u_{true} \end{bmatrix} = \begin{bmatrix} \dfrac{6P(L - d_{c,\sigma})}{WH^2} \\ \dfrac{2P{d_{c,u}}^2(3L - d_{c,u})}{EWH^3} \end{bmatrix} \tag{3.8}$$

W, H, and L are geometric parameters, as shown in Figure 3-3. P is the input load at the end of a cantilever beam and E is the elastic modulus. W, H, L, P, and E indicate the true parameters. $d_{c,\sigma}$ and $d_{c,u}$ are where the stress and displacement are measured. The subscripts 'σ' and 'u' represent stress and displacement, respectively. In this case study, the parameter L is considered to have an unrecognized error source, which means that L is incorrectly known. E and W are recognized unknown

parameters; thus, they need to be calibrated. Considering all error sources, the prediction model $\boldsymbol{\eta}$ is written as follows.

$$\boldsymbol{\eta} = \begin{bmatrix} \sigma_{\boldsymbol{\eta}} \\ u_{\boldsymbol{\eta}} \end{bmatrix} = \begin{bmatrix} \dfrac{6P(L_\delta - d_{c,\sigma})}{W_e H^2} \\ \dfrac{2P{d_{c,u}}^2(3L_\delta - d_{c,u})}{E_e W_e H^3} \end{bmatrix} \tag{3.9}$$

Here the prediction model σ and u involves the unrecognized sources of modeling error L_δ and unknown parameters E_e and W_e. Using (3.9), the D-optimality based experimental design method is performed to find the suitable $d_{c,\sigma}$ and $d_{c,u}$. The gradient matrix $\boldsymbol{\nabla_\theta \eta}$ is as follows.

$$\begin{aligned} \boldsymbol{\nabla_\theta \eta} &= \begin{bmatrix} \dfrac{\partial \sigma_{\boldsymbol{\eta}}}{\partial E} & \dfrac{\partial \sigma_{\boldsymbol{\eta}}}{\partial W} \\ \dfrac{\partial u_{\boldsymbol{\eta}}}{\partial E} & \dfrac{\partial u_{\boldsymbol{\eta}}}{\partial W} \end{bmatrix} \\ &= \begin{bmatrix} 0 & -\dfrac{6P(L_\delta - d_{c,\sigma})}{W_e^2 H^2} \\ -\dfrac{2P{d_{c,u}}^2(3L_\delta - d_{c,u})}{E_e^2 W_e H^3} & -\dfrac{2P{d_{c,u}}^2(3L_\delta - d_{c,u})}{E_e W_e^2 H^3} \end{bmatrix} \end{aligned} \tag{3.10}$$

Using (3.10), the D-optimality metric M is as follows.

$$\begin{aligned} M &= \det\left\{\left|(\boldsymbol{\nabla_\theta \eta})^{\mathrm{T}}(\boldsymbol{\nabla_\theta \eta})\right|\right\}^{-1} \\ &= \frac{E_e^2 W_e^3 H^5}{12P^2 d_{c,u}^2 (3L_\delta - d_{c,u})(L_\delta - d_{c,\sigma})} \end{aligned} \tag{3.11}$$

Using the parameter values shown in Table 3-1, the D-optimality metric **M** is calculated with regard to $d_{c,\sigma}$ and $d_{c,u}$, and shown in Figure 3-4. The result shows that **M** reaches its minimum at $d_{c,u} = L_\delta (=19)$ and $d_{c,\sigma} = 0$. The locations selected by the D-

optimality metric are identical to the locations of maximum stress and displacement, which is intuitively reasonable. The displacement and stress of the cantilever beam, in this example, change monotonically along the measurement location $\mathbf{D_c}$. With the monotonicity, the gradient $\boldsymbol{\nabla_\theta \eta}$ increases as the value of the response increases. For a highly nonlinear model, however, the maximum response does not always guarantee the D-optimality. This is further discussed in chapter 3.3.2, the engineering case study.

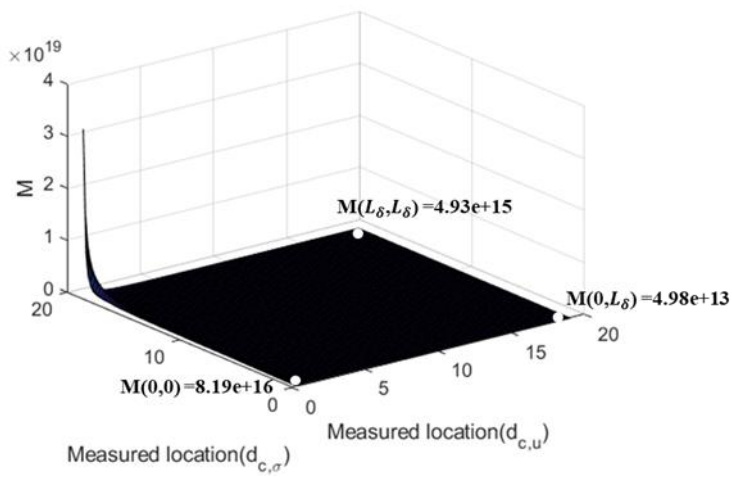


Figure 3-4 D-optimality metric (M) for the cantilever beam

Table 3-1 The model parameters and corresponding true values for the cantilever beam

Parameters	True values	Parameters	True values
E	200 (Gpa)	W	0.5 (m)
L	20 (m)	L_{δ}	19 (m)
H	0.5 (m)	P	15 (kN)
ε_{σ}	$0.1\max(\sigma_p)$	ε_u	$0.1\max(u_p)$

Next, these experimental points are used to calibrate the unknown parameters. The observation data $\boldsymbol{\eta}_o$ at the given experimental points has experimental error and is written as follows.

$$\boldsymbol{\zeta} = \begin{bmatrix} \sigma_{\boldsymbol{\zeta}} \\ u_{\boldsymbol{\zeta}} \end{bmatrix} = \begin{bmatrix} \dfrac{6P(L - d_{c,\sigma})}{WH^2} + \varepsilon_\sigma \\ \dfrac{2Pd_{c,u}^2(3L - d_{c,u})}{EWH^3} + \varepsilon_u \end{bmatrix} \tag{3.12}$$

where ε_σ and ε_u are the measurement error of σ and u at $d_{c,\sigma}$ and $d_{c,u}$. The calibration result is obtained by solving two equations, $\sigma_{\boldsymbol{\mu}} = \sigma_{\boldsymbol{\zeta}}$ and $u_{\boldsymbol{\mu}} = u_{\boldsymbol{\zeta}}$. The results of model calibration are shown in (3.13) and (2.14).

$$W_e = \frac{6P(L_\delta - d_{c,\sigma})W}{6P\left(L - d_{c,\sigma}\right) + WH^2\epsilon_\sigma} \tag{3.13}$$

$$E_e = \frac{Ed_{c,u}^2(3L_\delta - d_{c,u})\left\{6P\left(L - d_{c,\sigma}\right) + WH^2\epsilon_\sigma\right\}}{3(L_\delta - d_{c,\sigma})\left\{2Pd_{c,u}^2\left(3L - d_{c,u}\right) + \epsilon_u EWH^3\right\}} \tag{3.14}$$

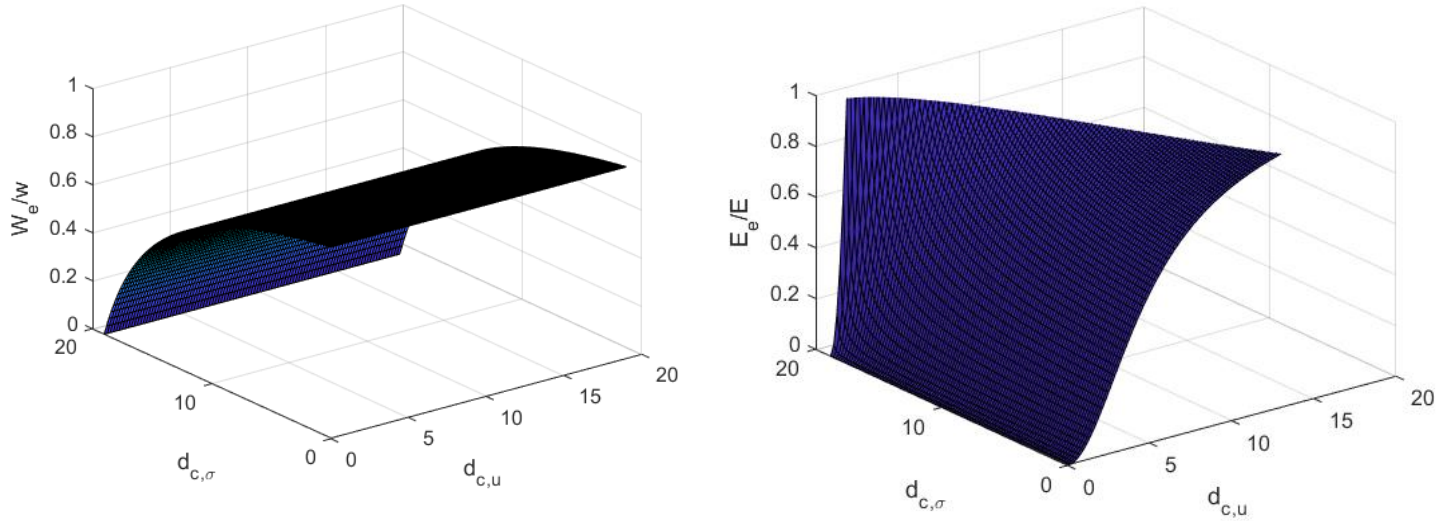
To compare the estimates and true parameters, (3.13) and (3.14) are normalized by the true parameters, resulting in (3.15) and (3.16). The estimates of the unknown parameters are accurate, if the normalized estimates W_e/W and E_e/E are close to one.

$$\frac{W_e}{W} = \frac{6P(L_\delta - d_{c,\sigma})}{6P\left(L - d_{c,\sigma}\right) + WH^2\epsilon_\sigma} \tag{3.15}$$

$$\frac{E_e}{E} = \frac{d_{c,u}^2(3L_\delta - d_{c,u})\left\{6P\left(L - d_{c,\sigma}\right) + WH^2\epsilon_\sigma\right\}}{3(L_\delta - d_{c,\sigma})\left\{2Pd_{c,u}^2\left(3L - d_{c,u}\right) + \epsilon_u EWH^3\right\}} \tag{3.16}$$

From (3.15) and (3.16), the conditions in which the normalized estimate is close

to one are achieved when the experimental noise ϵ_σ and ϵ_u are zero and the unrecognized uncertainty L_δ is close to L. The factor that is able to make the normalized estimate close to one is $d_{c,\sigma}$ and $d_{c,u}$. Building upon the previous discussion, if $d_{c,\sigma}$ and $d_{c,u}$ are close to the value designated by the D-optimality condition (e.g., $d_{c,\sigma}$ =0 and $d_{c,u} = L_\delta$) , the normalized estimate must be close to one. Figure 3-5 illustrates W_e/W and E_e/E with regard to $d_{c,\sigma}$ and $d_{c,u}$. Figure 3-5 (a) shows the normalized estimates with regard to $d_{c,\sigma}$ and $d_{c,u}$. In the figure, W_e/W depends only on $d_{c,\sigma}$, and is close to one when $d_{c,\sigma}$ =0, the D-optimal point. Similarly, the normalized estimate E_e/E is shown in Figure 3-5 (b); it depends on both $d_{c,\sigma}$ and $d_{c,u}$. In Figure 3-5 (b), the value of $d_{c,u}$, which enables W_e/W to be one, is different according to the value of $d_{c,\sigma}$. When $d_{c,u}$ =0, the value of $d_{c,u} = L_\delta$ enables W_e/W to approach one. These results support that D-optimality based experimental design reduces the estimation error in model calibration.



(a) (b)

Figure 3-5 Estimation errors of unknown parameters (a) width (W_e/W), (b) Young's modulus (E_e/E)

For efficient model refinement activities, the invalidity of model prediction should be that which arises mainly due to unrecognized sources of modeling errors. This research compares the errors of predicted responses using the estimates of unknown parameters and the true value of the unknown parameters. Figure 3-6 describes the model prediction errors due to model form errors and parameter estimation errors. The model prediction with model form errors are plotted by the black dotted line. The model prediction with model form and parameter estimation errors are plotted by the orange line. The discrepancy between two plots are the prediction errors due to parameter estimation errors. Figure 3-6(a) illustrates the prediction errors of stress response and Figure 3-6(b) illustrates the prediction errors of displacement response. The graphs in Figure 3-6 show that the prediction errors due to parameter estimation errors decreases when the measurement point reaches D-optimality based experimental design (e.g., $d_{c,u}=L_\delta$ and $d_{c,\sigma}=0$). However, Figure 3-6(b) shows the limitation that the discrepancy increases when the measurement location of displacement response reaches at $d_{c,u}=L_\delta$. It is due to the effect of model form errors in experimental design. Since the experimental design process is performed with the initial prediction model which includes model form errors, the result of experimental errors cannot ensure the least parameter estimation errors. To solve this problem, the model calibration process should consider the effect of model form errors in prediction. Chapter 4 deals with this issue.

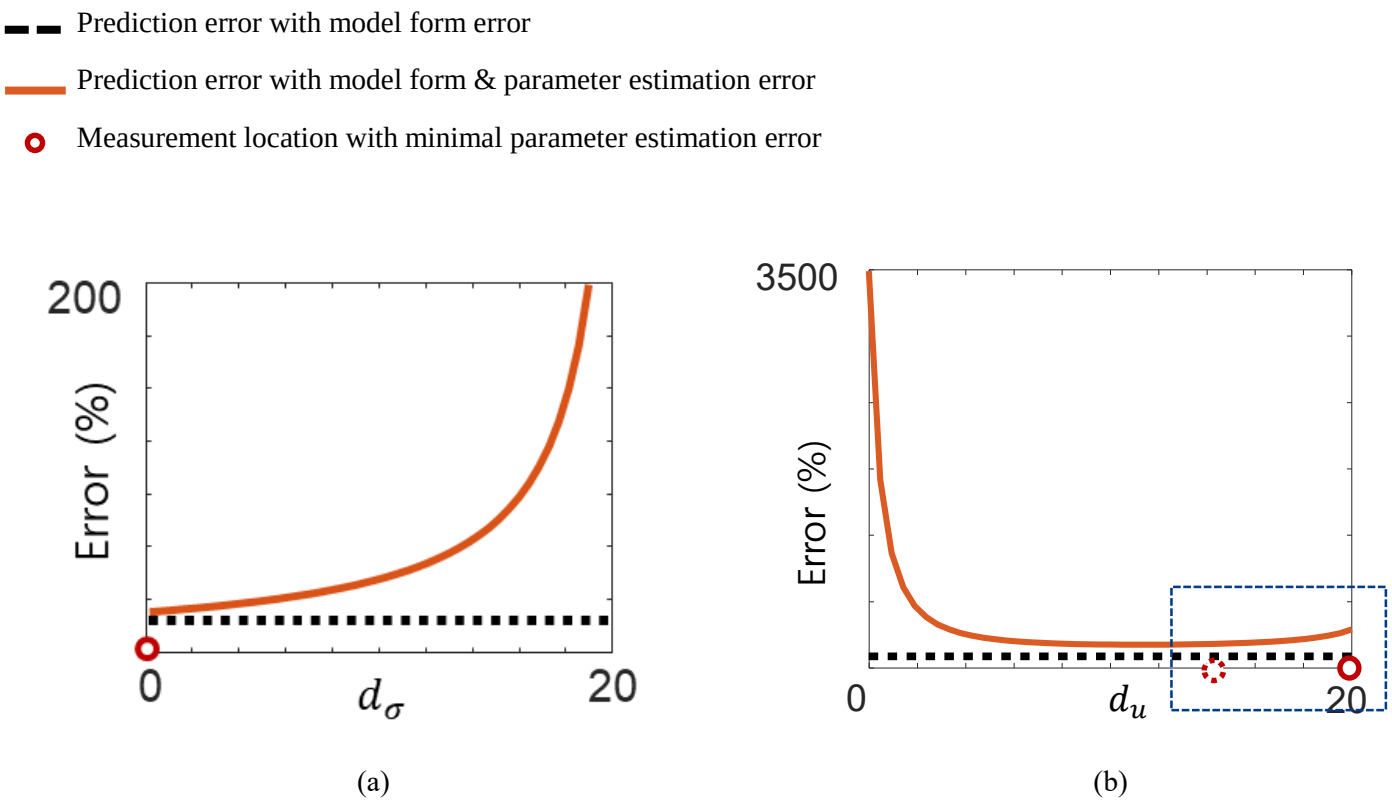
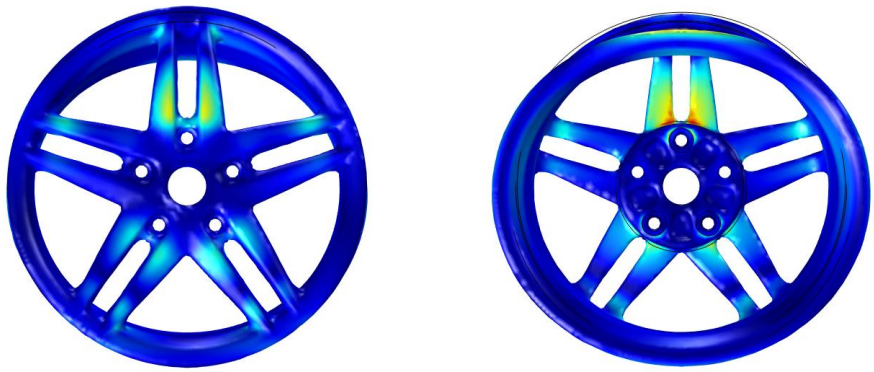


Figure 3-6 The prediction errors due to model form and estimation of unknown parameters; (a) the stress prediction, (b) the displacement prediction

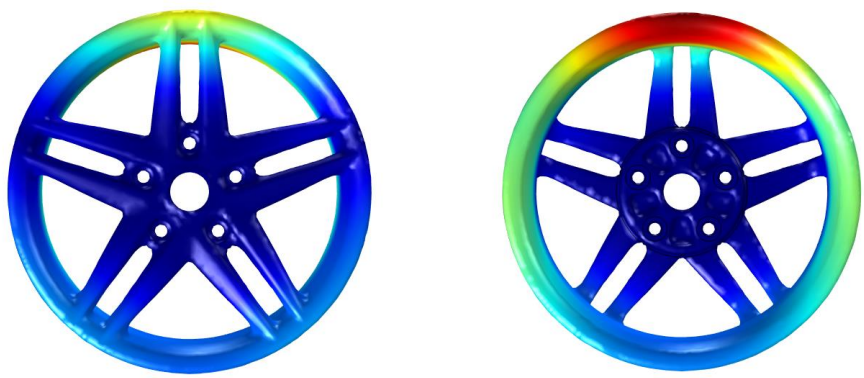
3.3.2 Engineering Case Study: Automotive Wheel Rim FEM Model

An automotive wheel rim is the outer frame of a wheel in an automobile that holds a tire. The model of the rim used in this study is provided by Application Libraries in COMSOL software (Multiphysics 1998) (Tabatabaian 2015). The structural analysis computes the stress and displacement of a wheel against the automobile weight and the tire pressure. Figure 3-7 and Figure 3-8 show the stress and displacement analysis result of the wheel rim model. Figure 3-9 provides all given loads of the automotive wheel rim model. Figure 3-9 (a) denotes the fixed boundary condition that represents that the inner frame of the rim fixed to an automobile body frame. Figure 3-9 (b) and (c) describe the distributed load from the tire pressure and weight of the automobile. The wheel rim model includes modeling error in the fixed boundary condition for unrecognized error sources. The five fixed boundary locations, as shown in Figure 3-9 (a) are reduced to three fixed locations in Figure 3-10. The boundary conditions in Figure 3-10 are adopted in the prediction model. The model with a fixed boundary condition in Figure 3-9 (a) provides the observation data. Furthermore, the observation includes experimental error $\boldsymbol{\varepsilon}= 0.001\zeta$, which denotes the 0.001 times of each observation. The main parameters, including the material properties, are described in Table 3-2. The wheel rim is made of aluminum. For model calibration, this study designates two unknown parameters; the elastic modulus of aluminum and the pressure of the tire.



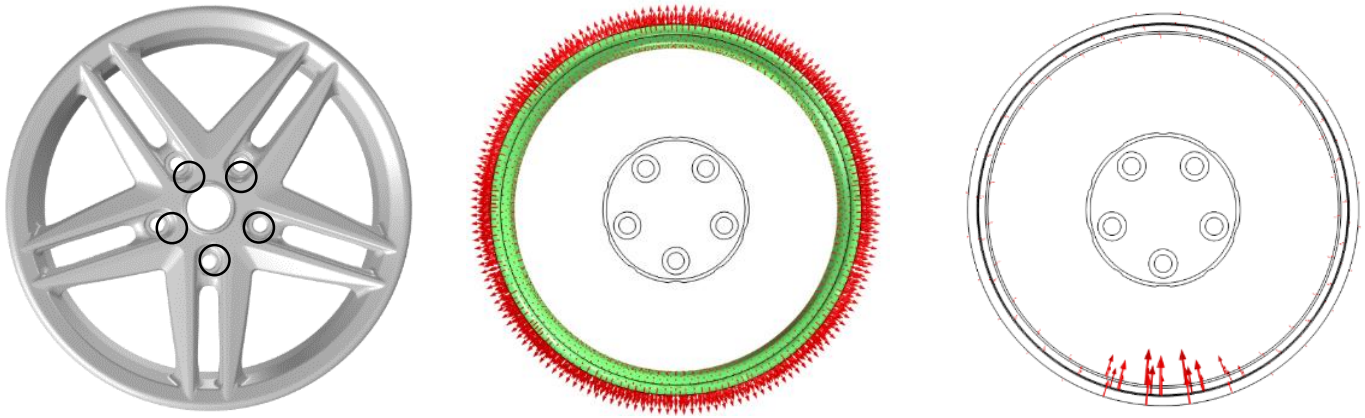
(a) (b)

Figure 3-7 Stress analysis of the wheel rim model; (a) front view, (b) rear view



(a) (b)

Figure 3-8 Displacement analysis of the wheel rim model; (a) front view, (b) rear view



(a) (b) (c)

Figure 3-9 Load conditions of an automotive wheel rim model; (a) fixed locations (five locations noted with black circles), (b) distributed load from the tire pressure, (c) distributed load from the weight of the automobile

Table 3-2 The model parameters and corresponding true values for the cantilever beam

Parameters	True values	Known / Unknown
Elastic modulus (E)	70.00 [GPa]	Unknown
Poission's ratio (ρ)	0.330	Known
Pressure of the tire ($Pr.$)	2.000 [bar]	Unknown
Weight of an automobile (W)	1120 [kg]	Known

Figure 3-10 Fixed locations for unrecognized sources of modeling error

Figure 3-11 presents 70 observable points used by D-optimality based experimental design to select a combination of observations. 65 points are located in the front-view in Figure 3-11 (a); five points are located in the rear-view in Figure 3-11 (b). The points are concentrated in half of the wheel rim model, since the geometry and load condition are symmetric. At each point, two different responses (e.g., stress and displacement) are acquired. Thus, the overall number of observations is 140. This generates ${}_{140}C_2$ combinations of observations. Among these combinations, an observation set that gives a minimum M metric value, mentioned in (3.5), becomes the D-optimality based experimental design for model calibration.

(a) (b)

Figure 3-11 Observable candidate locations for the wheel rim model (a) front view, (b) rear view

For M metric calculation, this study adopted a finite difference method to easily calculate the term $\nabla_{\theta}\eta$ at E = 68GPa and $Pr.$=1.8bar, the initial values for model calibration. Figure 3-12 shows the D-optimality based experimental design (M). The combination of observations in Figure 3-12 gives the minimal M = 2.57e-08. Comparing Figure 3-12 with Figure 3-7 and Figure 3-8, the selected locations are not where the maximum stress and displacement occur. This is due to the geometric nonlinearity of the wheel rim model, which results in a nonlinear prediction about the observable location. The results of the model calibration performed with both the proposed method and with the maximum responses are compared in Figure 3-13 to show the efficacy of D-optimality based experimental design for model improvement.

(a) (b)

Figure 3-12 D-optimal design of observations (a) location for stress, (b) location for displacement

(a) (b)

Figure 3-13 The experimental design that gives the maximal responses; (a) stress (front view), (b) displacement (rear view)

Table 3-3, Table 3-4, and Table 3-5 summarize the model calibration results using D-optimal design and maximum responses design, as shown in Figure 3-12 and Figure 3-13. Table 3-3 and Table 3-4 include the estimates of unknown parameters and normalized estimates E_e/E and Pr_e/Pr. The estimates of unknown parameters are accurate, as the normalized estimates are close to one. Examining the results in Table 3-3, Table 3-4, it can be seen that the estimates of unknown parameters determined by D-optimal design are much closer to the true values than are the results of maximal response design. The result in Table 3-5 gives the predicted response errors using the estimates of unknown parameters and the true value of the unknown parameters. Table 3-5 uses the normalized root mean squared errors (RMSE), as shown in (3-20).

$$NRMSE = \frac{N}{\sum_{i=1}^{N} \boldsymbol{\eta}(\boldsymbol{\theta}, \boldsymbol{x}; \boldsymbol{D_i})} \sqrt{\frac{1}{N} \sum_{i=1}^{N} \{\boldsymbol{\eta}(\boldsymbol{\theta_c}, \boldsymbol{x}, \boldsymbol{\delta}; \mathbf{D}_i) - \boldsymbol{\eta}(\boldsymbol{\theta}, \boldsymbol{x}, \boldsymbol{\delta}; \mathbf{D}_i)\}^2} \quad (3.17)$$

The terminology $\boldsymbol{\theta}$, $\boldsymbol{\theta_c}$, $\boldsymbol{x}$, $\boldsymbol{\delta}$, and $\mathbf{D}$ is the same as that used in (3.1). N is the number of observable points for each predicted response, displacement and stress. The predicted response errors in (3.20) are due to parameter estimation errors in model calibration. The result of D-optimal design shows the smaller normalized RMSE than the maximal responses design in Figure 3-13. Overall, the results in Table 3-5 show that the D-optimality based experimental design gives a prediction that is closer to the prediction that only includes modeling errors.

Table 3-3 Comparison of the model calibration results of unknown variable E (Young's modulus)

	D-optimal design	Maximal response design
Estimated E (Gpa)	65.14	65.00
E_e/E	0.9305	0.9009

Table 3-4 Comparison of the model calibration results of unknown variable P (pressure of a tire)

	D-optimal design	Maximal response design
Estimated P (bar)	1.802	1.673
Pr_e/Pr	0.9285	0.8365

Table 3-5 Comparison of the normalized RMSE of the predicted responses (stress and displacement)

	D-optimal design	Maximal response design
Stress	1.881E-4	2.114E-4
Displacement	5.668E-3	6.030E-3

By the parameter estimation result of combination 2 in Table 3-3 and Table 3-4, the estimation accuracy of E is higher than that of Pr. The difference originates from the sensitivity of the responses about the parameters. Since the selected responses combination in combination 2 are more sensitive to E, The optimization algorithm is prone to easily change the value of E. Table 3-6 describes the sensitivity of each selected responses in combination 1 and combination 2 about the model parameters E and Pr. Table 3-6 describes the sensitivity about the model parameters E and Pr. In this table, the terms $\frac{\partial \eta_i}{\partial E}$ and $\frac{\partial \eta_j}{\partial P}$ each denote the sensitivity term about E and Pr. The numbers in brackets stand for the rank of the sensitivity about each parameters in all 140 responses. For example, the sensitivity of the stress about E in combination 1 is 2.5880E-05. The value is the 29th largest among 150 responses. The third row of the table includes the summation of normalized sensitivity, calculated as the following equation.

$$\omega_j = \frac{\frac{\partial \eta_j}{\partial E}}{\sum_{i=1}^{140} \frac{\partial \eta_i}{\partial E}} + \frac{\frac{\partial \eta_j}{\partial P}}{\sum_{i=1}^{140} \frac{\partial \eta_i}{\partial P}} \tag{3.18}$$

By (3.21), the sensitivity about the parameter E and Pr is easily integrated with normalized sensitivity term. The result in Table 3-6 shows that the responses in combination 1 has the first and second largest normalized sensitivity. The result implies that the parameter estimation accuracy depends on the sensitivity of the responses, not the magnitude. However, the responses combination which only gives the highest sensitivity about the unknown parameters can lead unidentifiable estimation of the parameters. For example, the response combination gives a parameter solution as a continuous region, not a discrete value.

Table 3-6 Sensitivity analysis of each response in combination 1 and 2

	Combination 1 (Proposed-determinant)		**Combination 2 (Maximum response)**	
	η_1(Stress)	η_2 (Displacement)	η_3 (Stress)	η_4 (Displacement)
$\frac{\partial \eta_i}{\partial E}$	2.5880E-05 (29/140)	1.9272E-04 (1/140)	3.6595E-06 (83/140)	1.9272E-04 (1/140)
$\frac{\partial \eta_j}{\partial P}$	0.8320 (1/140)	6.6119E-06 (74/140)	0.0385 (61/140)	6.6119E-06 (74/140)
ω_j	0.0616(2/140)	0.0754 (1/140)	0.0038 (97/140)	0.0754(1/140)

3.4 Summary and Discussion

This chapter proposed a method for optimization-based model improvement with experimental design to efficiently handle the error sources of an invalid computational model. Existing optimization-based model improvements have a limitation in that the error sources that affect their validity are indistinguishable. To overcome this, D-optimality based experimental design method is integrated to identify the major invalidity sources of a computational model by reducing the estimation errors in the model calibration. By doing so, the model validation step can evaluate the invalidity of a computational model due to unrecognized errors, and the model refinement step can efficiently examine the main cause of invalidity from among the various candidates of unrecognized modeling errors.

The benefits of the proposed method are demonstrated by examining two case studies. A numerical example of a cantilever beam shows the feasibility of the D-optimality based experimental design in the model improvement process. The study shows that the parameter estimation errors in model calibration are reduced when the measurement location becomes close to the D-optimality-based experimental design. Another engineering example (e.g., an automobile wheel rim) is examined to show the application of the proposed method in practice. Based on the estimates of unknown parameters using D-optimality experimental design, the overall amount of errors in the predicted responses are close to the amount of errors that arise due to an invalid modeling. The identified modeling errors in a computational model are then evaluated using a model validation process and the validity of the model can be improved by the model refinement step.

The proposed method can be widely applied where computational models are used, such as for study of a digital twin in a cyber-physical system (CPS). Since a CPS is composed of a large number of computational models, it is important to find an effective way to construct models, in terms of development time and budget. The proposed method is a preliminary study to reduce the modeling cost required to ensure the accuracy of digital twin models. Furthermore, enhanced estimates of unknown parameters in an experimental design can extend the valid domain of the computational model. This method can also help with parameter updating during the operation of the computational model to synchronize with reality.

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

1) **Hyejeong Son**, Byeng D. Youn, and Taejin Kim, "Model Improvement with Experimental Design for Identifying Error Sources in a Computational Model," *Structural and Multidisciplinary Optimization*, Accepted, 2021.

Chapter 4

Proportionate Bias Calibration with Bound Information to Consider Unrecognized Model Form Errors

Model calibration is a process of estimating unknown model parameters as a tool of improving a computational model (Park and Himmelblau 1982) (Kennedy and O'Hagan 2001) (Oberkampf and Roy 2010) (Thonhofer et al. 2014) (Qiu et al. 2018). It is treated as an inverse problem solving an n-th equation which indicates that the predicted response is the same as the observation data. Optimization-based model calibration (OBMC) is one of the popular techniques of model calibration, which employs optimization algorithms in a deterministic manner (Frank and Shubin 1992) (Hills and Trucano 2002) (Gholizadeh 2013) (Lee et al. 2019). In OBMC, the optimization method allows nonlinear predicted responses, which has difficulty finding an analytical solution to get an approximate result.

Despite these advantages, the model calibration suffers from parameter estimation errors due to a variety of error sources. As stated in Chapter 3, the coupled error sources between the parameter estimation errors and model from errors disrupt

model validation to make the right decision for a computational model. Thus, it is of importance to reduce the parameter estimation errors to relieve the imprecise decision in model validation. Although the integration of experimental design proposed in Chapter 3 can alleviate the problem, the computational model after the model calibration is still troubled with parameter estimation errors. The parameter estimation errors are affected by the model form errors and measurement errors that is hard to discover. In practical fields, engineers can doubt the estimated unknown parameter by the OBMC because some estimated results are different from their experience-based information, such as an approximate bound or dimension of the value of unknown parameters.

To solve this problem, the related scholars have tried to quantify model form errors and consider the effect of model form errors while the model calibration. One approach is Bayesian-based model calibration with bias correction that is applicable with observation data measured in the diverse design domains (Kennedy and O'Hagan 2001) (McFarland et al. 2007) (Higdon et al. 2008) (Arendt et al. 2010) (Plumlee 2017). The word 'bias' in this method refers to the Gaussian process model for the amount of model form errors, acquired with the prior information about the observation data and initial predicted responses. When the unknown model parameter is updated by the Bayes theorem, the bias term helps the model calibration to estimate a reasonable value for unknown model parameters.

This chapter devices proportionate bias calibration with prior bound information to consider the effect of model form errors in OBMC. In this method, the bias term is treated as another unknown parameter for optimization. The bound information of unknown parameters can play a crucial role in OBMC that it is one

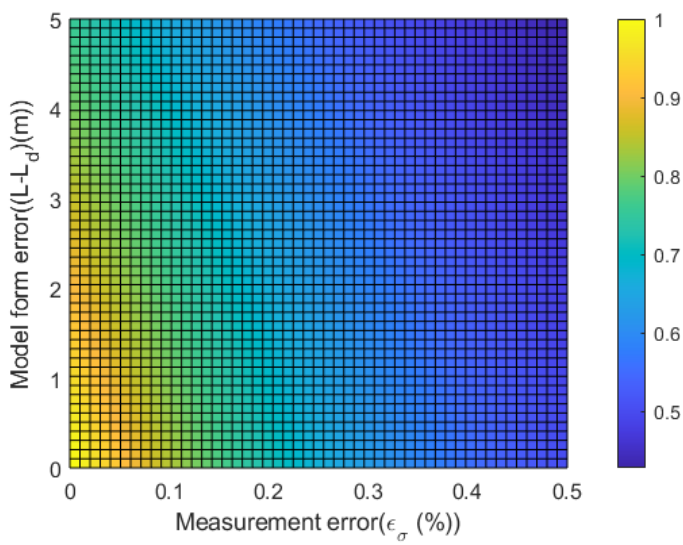
of the constraints in the optimization algorithm. Besides, this chapter devices output-dependent bias, a new formulation for model form errors, to support the optimization with multiple responses which dimensions are severely different.

The remainder of Chapter 4 is organized as follows. Chapter 4.1 presents brief review of optimization-based model calibration step which suffers from the effect of model form errors. Chapter 4.2 provides the detailed descriptions of the proposed method, the definition of proportionate bias and model calibration with bound information. Chapter 4.3 discusses the efficiency of the proposed method with analytical and engineering case studies, as introduced in Chapter 3.3.

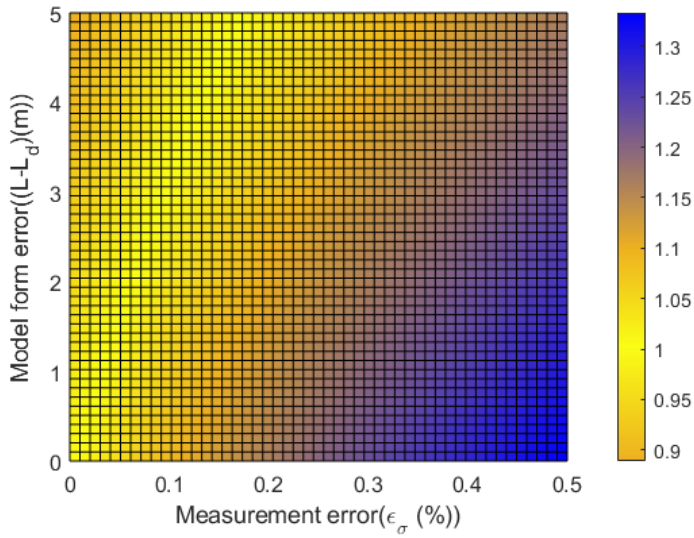
4.1 Limitations of Experimental Design for OBMI with the Effect of Model Form Errors

Chapter 3 proposed a new experimental design for efficient OBMI, which selects the optimal experimental design to reduce the parameter estimation errors by minimizing the specific term in (3.7). The method is efficient when reducing the parameter estimation errors without additional information. However, it remains a limitation to remove all amount of parameter estimation errors. Equation (3.7) consists of the following terms: $\nabla_{\theta}\boldsymbol{\eta}$, $\boldsymbol{\varepsilon}$, and $\boldsymbol{\delta}$. The proposed experimental design aims to minimize the term $\nabla_{\theta}\boldsymbol{\eta}$, since the other two terms $\boldsymbol{\varepsilon}$, and $\boldsymbol{\delta}$ are uncontrollable. By the equation (3.7), the parameter estimation errors are affected by the model form errors and measurement errors as well as the gradient of the prediction $\nabla_{\theta}\boldsymbol{\eta}$. Depending on the severity of error sources, the value of estimates for unknown parameters might be against the basic knowledge about the model parameters.

Figure 4-1 shows the effect of error sources on the parameter estimation errors with the cantilever beam example, as discussed in Chapter 3.3.1. The figures show the parameter estimation errors derived in (3.15) and (3.16) depending on two error sources (e.g., measurement and model form errors) when the observation is measured at the optimal experimental design. Figure 4-1 (a) and (b) respectively illustrates the estimation errors of width (W) and young's modulus (E). When the value is close to one, the surface color in Figure 4-1 is close to yellow. According to (3.15) and (3.16), the closer the value is to one, the closer the estimates are to the true value. Figure 4-1 (a) presents that the estimation error of W is reduced to 50% of the actual value, even though the model calibration utilized the optimal design for observation data. Figure 4-1 (b) gives the estimation error of E differs from the actual value by more than about 30%.



(a)



(b)

Figure 4-1 The parameter estimation errors of the cantilever beam with optimal experimental design (a) Width (Wp/Wt), (b) Young's modulus (Ep/Et)

The result in Figure 4-1shows that the experimental design selection encounters a limitation to effectively improving the parameter estimation errors when severe errors exist. This fact gives a doubt on the effectiveness of OBMI since the estimated values of unknown parameters are against the prior information about the prediction models that include an approximate range or statistical information of model parameters. Thus, the proposed method in this chapter aims to consider the amount of error sources in the process of model calibration to draw reasonable estimates for unknown parameters.

4.2 Proportionate Bias Calibration with Bound Information of Model Parameters

This chapter aims to perform the model calibration with consideration of the error sources in a prediction model. The bound information of unknown model parameters are adopted as a constraint in OBMC to roughly quantify the errors. Chapter 4.2.1 presents the definition of proportionate bias. Chapter 4.2.2 explains the framework of proportionate bias calibration with bound information.

4.2.1 The Formulation of Proportionate Bias

For the consideration of error sources, model calibration adopts the formulation of prediction and observation with the discrepancy function (Arendt et al. 2010) (Reichert and Schuwirth 2012) (Li et al. 2016) (Qiu et al. 2018). As a previous study, the bias correction is the one of popular method, generally used with Bayesian-based model calibration. The formulation of model calibration with bias function is as follows.

$$\boldsymbol{\delta}(\boldsymbol{\theta}, \boldsymbol{x}) + \boldsymbol{\eta}(\boldsymbol{\theta}, \boldsymbol{x}) = \boldsymbol{\zeta}(\boldsymbol{x}) + \boldsymbol{\varepsilon} \tag{4.1}$$

The discrepancy $\boldsymbol{\delta}$ is dependent to unknown parameters ($\boldsymbol{\theta}$) and design parameters ($\boldsymbol{x}$). The method quantifies Gaussian process model for the discrepancy, which involves model form errors and measurement errors, through the observation and prediction data along the model parameters. As stated in Chapter 2, the method is efficient when the observation data measured in various design sites and reliable prior information are available.

When the observation data only measured in few number of design sites, Qiu et al. has devised sensitivity-based parameter calibration for consideration of model form errors (Qiu et al. 2018). The formulation of model calibration is as follows.

$$\boldsymbol{\zeta}(\boldsymbol{x}) = \boldsymbol{\eta}(\boldsymbol{\theta}, \boldsymbol{x}) + \delta \tag{4.2}$$

δ denotes the sensitivity-based parameter in (4.2). Figure 4-2 shows the process of sensitivity-based parameter calibration. In this figure, $\mathbf{X}_{\mathrm{k}}$ denotes a vector of known parameters and $\mathbf{X}_{\mathrm{unk}}$ denotes a vector of unknown model parameters. The method assumes that the discrepancy is a constant function, the simplest form of regression. This assumption enables model calibration with the observation data measured at three design sites. The sensitivity-based model calibration firstly estimates the optimal δ, matching the sensitivity of prediction and observation the same. With the optimal δ, optimal values of unknown model parameters are estimated.

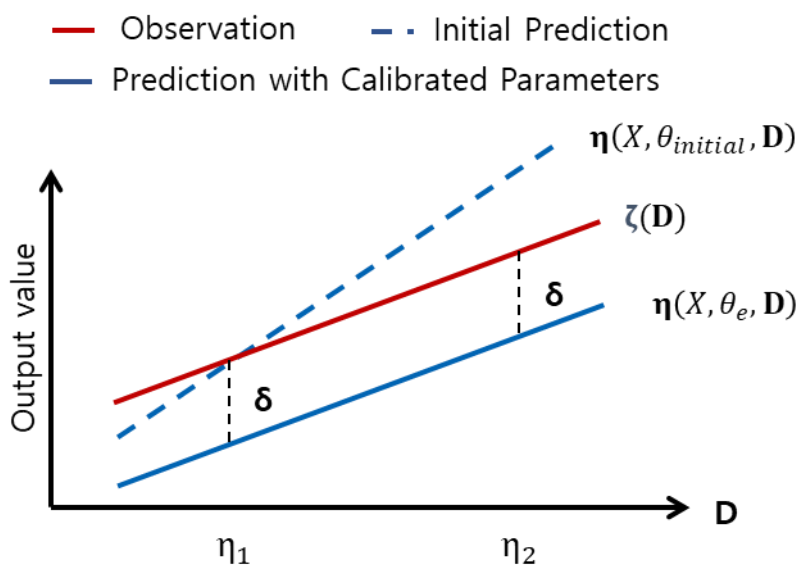


Figure 4-2 Sensitivity-based parameter calibration

In the sensitivity-based parameter calibration, the constant function for the discrepancy assumption is valid, especially when the magnitude of output values are similar. If the magnitude of the responses is highly different, model calibration optimizes the constant discrepancy by adjusting only a specific response among all responses, leading to parameter estimation errors. This chapter devises a new formulation for the discrepancy, called proportionate bias, to solve this problem. The proportionate bias assumes that the formulation of the discrepancy is a ratio of output value. The model calibration formulation with the proportionate bias is as follows.

$$\boldsymbol{\zeta}(\boldsymbol{x}) = \delta \times \boldsymbol{\eta}(\boldsymbol{\theta}, \boldsymbol{x}) \tag{4.3}$$

Equation (4.3) defines the observation as the multiplication of the proportionate bias δ and the prediction response $\boldsymbol{\eta}(\boldsymbol{\theta}, \boldsymbol{x})$. Figure 4-3 explains the process of the proportionate bias calibration. This figure depicts that the proportionate bias is estimated to 0.5 for η_1, η_2 and η_3. The OBMC estimates the suitable value of proportionate bias with the unknown model parameters. The bound information of unknown parameters determines the lower and upper bound of the proportionate bias for the constraint in OBMC. Chapter 4.2.2 discusses the details of proportionate bias calibration.

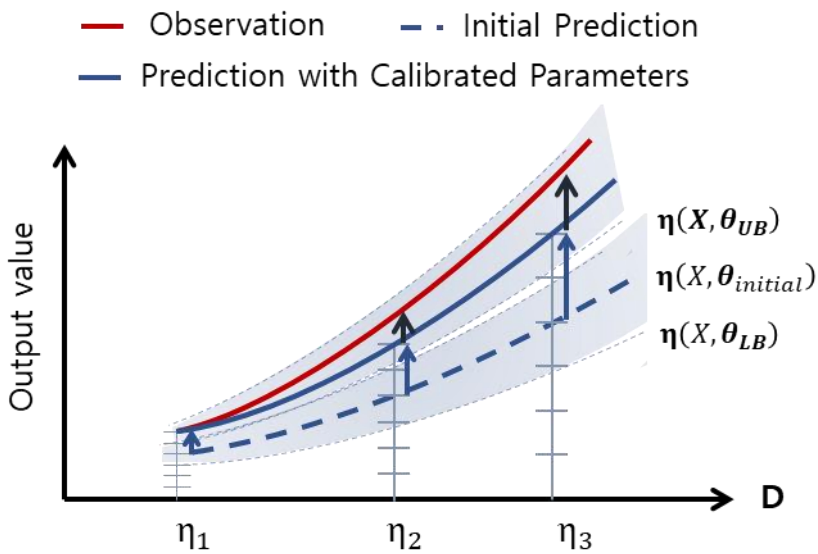


Figure 4-3 The proportionate bias calibration

4.2.2 Proportionate Bias Calibration with Bound Information of Unknown Model Parameters

The objective of proportionate bias calibration is to increase the accuracy of estimated unknown model parameters by considering the error sources when calibrating unknown model parameters. In this method, the bound information of unknown model parameters can provide a guide to estimate a reasonable value for unknown parameters and the discrepancy. Figure 4-4 shows the overall process of proportionate bias calibration using the bound information. Comparing with the Figure 2-1 explaining the overall framework of OBMC, the proportionate bias calibration adopts the bound information of unknown parameters and proportionate bias. Furthermore, the proportionate bias (δ) is optimized with unknown model parameters.

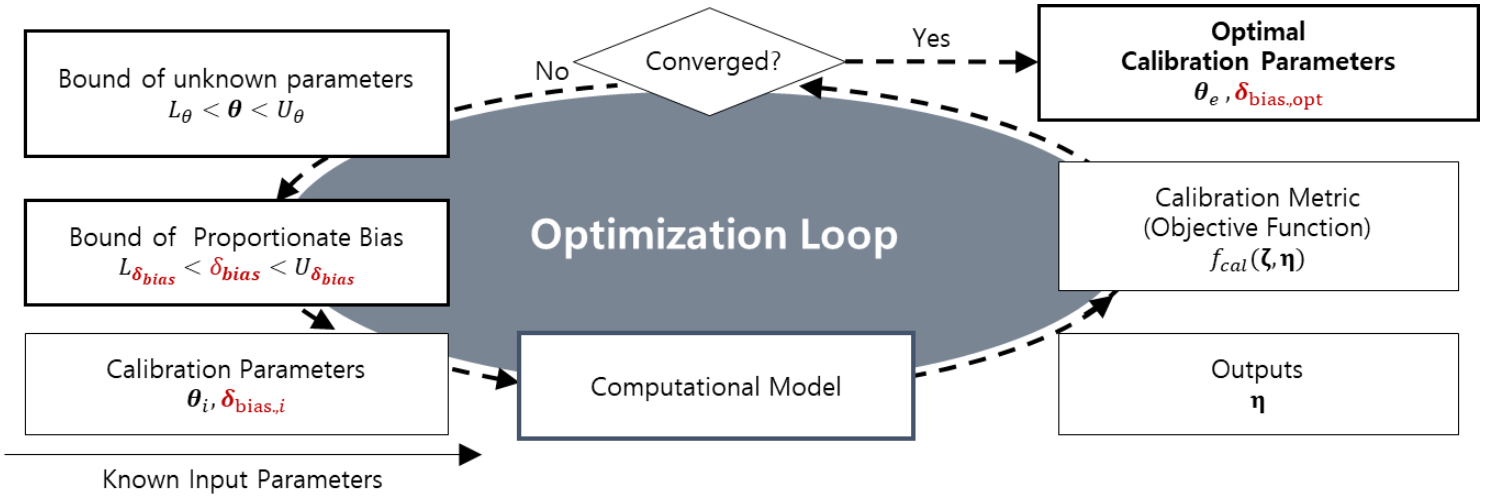


Figure 4-4 The proportionate bias calibration with bound information

In this process, the bound of unknown parameters can determine the bound of proportionate bias. (4.4) defines a bound vector of unknown parameter $\boldsymbol{\theta}_B$, which consists of upper or lower bound of each model parameter.

$$\boldsymbol{\theta}_{B_k} = \left[\theta_1^j, \theta_2^j, \cdots, \theta_n^j\right] \ \ where \ \ j = [\mathrm{LB}, \mathrm{UB}] \tag{4.4}$$

n refers to the number of unknown model parameters. j designates the upper or lower bound. LB and UB denotes the lower bound and upper bound. Since each element can be upper or lower bound of the parameters, the number of possible $\boldsymbol{\theta}_B$ is 2^n (e.g., k=1, $\cdots 2^n$). The following equation represents the upper and lower bound of the proportionate bias according to the bound of unknown parameters.

$$\mathbf{L}_\delta = \min_k \frac{\boldsymbol{\zeta}(\boldsymbol{x})}{\boldsymbol{\eta}\left(\boldsymbol{\theta}_{\mathrm{B}_k}, \boldsymbol{x}\right)} \tag{4.5}$$

$$\mathbf{U}_\delta = \max_k \frac{\boldsymbol{\zeta}(\boldsymbol{x})}{\boldsymbol{\eta}\left(\boldsymbol{\theta}_{\mathrm{B}_k}, \boldsymbol{x}\right)} \tag{4.6}$$

4.3 Case Studies

This chapter adopts the analytical and engineering models for the case studies, already explained in Chapter 3.3.1 and 3.3.2. With three case studies, the detailed process of proportionate bias calibration is provided. To verify the efficacy of the proposed method, the case studies compare the estimates of unknown model parameters calibrated through the proposed method and the sensitivity-based parameter calibration.

4.3.1 Analytical Case Study: Cantilever Beam Model

This chapter applies the proportionate bias calibration to the cantilever beam model, same as Figure 3-3. The model calibration is performed by four calibration approach; Bayesian method, OBMC with bound information, OBMC with sensitivity bias, and OBMC with proportionate bias. The model includes model parameters, model form errors, and experimental errors, as shown in Table 3-1. The model calibration selected the observation data as displacement responses measured at $d_{c,u}$ =0.5 L_δ and $d_{c,u}$ = L_δ . The unknown parameters are young's modulus (E) and width (W). This study assumes that the reasonable bound information of unknown parameters can be determined through the standard deviation of the parameters. The bound information of unknown parameters is set as [-3σ, 3σ], where σ denotes the standard deviation of the unknown parameters. The standard deviation of young's modulus and width is respectively 15.2Gpa and 0.0220m (Hess et al. 2002). With the statistical information, the bound of young's modulus and width are [154.4, 245.6] Gpa and [0.434, 0.566] m. The bound of proportionate bias is calculated using (4.5) and (4.6). Table 4-1 summarizes the process of bound selection for sensitivity-based parameter and proportionate bias. For Bayesian-based method, biased term is quantified by the Gaussian process (GP) modeling with the discrepancy of experimental data and initial prediction at at $d_{c,u}$ =0.5 L_δ and $d_{c,u}$ = L_δ . Since only two datasets are used for GP model of the discrepancy, GP model can give fitting errors to predict the amount of errors due to model form and measurement.

Table 4-1 The bound selection for sensitivity-based and proportionate bias

	E	**W**	**Prediction**	**Sensitivity-based**	**Proportionate bias**
σ	LB (154.4)	LB (0.434)	1.58E+07	-0.14E+07	0.9114
	LB (154.4)	UB (0.566)	1.21E+07	0.23E+07	1.1901
	UB (245.6)	LB (0.434)	1.58E+07	-0.14E+07	0.9114
	UB (245.6)	UB (0.566)	1.21E+07	0.23E+07	1.1901

	E	W	Prediction	Sensitivity-based	Proportionate bias
u	LB (154.4)	LB (0.434)	0.049	-0.0106	0.784
	LB (154.4)	UB (0.566)	0.038	0	1.000
	UB (245.6)	LB (0.434)	0.031	0.0074	1.239
	UB (245.6)	UB (0.566)	0.024	0.0144	1.600
	Lower Bound			**-0.14E+07**	**0.784**
	Upper Bound			**0.23E+07**	**1.600**

For the OBMC, this case study adopted the calibration metric which considers the discrepancy formulation. (4.7) is the calibration metric without considering the discrepancy function. (4.8) and (4.9) is the calibration metric for sensitivity-based parameter calibration and proportionate bias calibration, respectively.

$$\min_{\boldsymbol{\theta},\boldsymbol{\delta}} \sum_{i=1}^{n} \frac{\mathbf{1}}{\boldsymbol{\zeta_i}} \left(\boldsymbol{\zeta_i} - \boldsymbol{\eta_i}(\boldsymbol{X}, \boldsymbol{\theta}(E,\, W))\right)^2 \tag{4.7}$$

$$\min_{\boldsymbol{\theta},\boldsymbol{\delta}} \sum_{i=1}^{n} \frac{\mathbf{1}}{\boldsymbol{\zeta_i}} (\boldsymbol{\zeta_i} - \boldsymbol{\eta_i}(\boldsymbol{X}, \boldsymbol{\theta}(E,\, W)) - \boldsymbol{\delta})^2 \tag{4.8}$$

$$\min_{\boldsymbol{\theta},\boldsymbol{\delta}} \sum_{i=1}^{n} \frac{\mathbf{1}}{\boldsymbol{\zeta_i}} \left(\boldsymbol{\zeta_i} - \boldsymbol{\delta} \times \boldsymbol{\eta_i}(\boldsymbol{X}, \boldsymbol{\theta}(E,\, W))\right)^2 \tag{4.9}$$

Using (4.6), (4.7), and (4.8), Table 4-2 gives the result of model calibration. (4.9) quantifies the parameter estimation errors εrr_E^* and εrr_W^*.

$$\varepsilon rr_{\theta_i}^* = \frac{\left|\theta_{i,true} - \theta_{i,calibrated}\right|}{\theta_{i,true}} \tag{4.10}$$

Table 4-2 provides the optimized unknown parameters, the parameter estimation errors, and optimized discrepancy(δ). Figure 4-5 illustrates the parameter estimation errors of each four calibration method. Bayesian-based method gives the largest estimation errors, due to the errors in discrepancy function modeled by GP. OBMC with bound information shows the result which is close to the bound of the parameters. It implies that the optimal result is not exist in the given bound of the model parameters. The proportionate bias method gives the lowest parameter estimation errors in both two unknown parameters. The sensitivity-based method shows the most significant estimation errors since the invalid optimal value of the sensitivity-based model bias term is valid only for the displacement. For the stress response, the value '-0.0102' is too small to support the accuracy of the responses whose magnitude is 10^7 times larger than the optimal bias.

Table 4-2 The result of model calibration for the cantilever beam model

	Bayesian-based method	**OBMC**	**OBMC with Sensitivity bias**	**OBMC with Proportionate bias**
δ	- (GP)	-	-0.0102	1.0175
E(Gpa)	180.5	180.5	245.6	**199.65**
εrr_E^*	0.0975	0.0975	0.2280	**0.0017**
W (m)	0.4750	0.4750	0.4750	**0.4833**
εrr_W^*	0.0500	0.0500	0.0500	**0.0334**

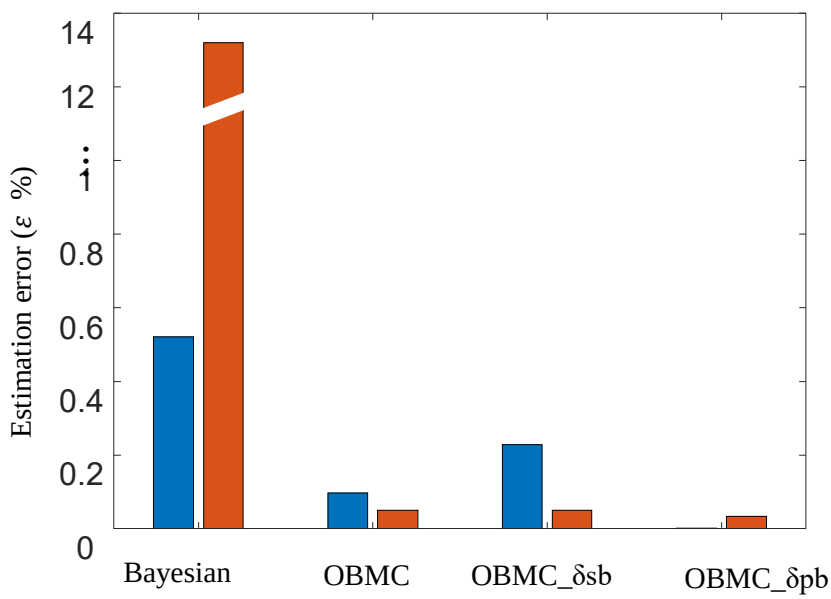


Figure 4-5 The proportionate bias calibration with bound information

4.3.2 Engineering Case Study 1: Automotive Wheel Rim FEM Model

Chapter 4.3.2 employs an engineering case study, already introduced in Chapter 3.3.2. The model includes model parameters, and experimental errors, as shown in Table 3-2. The considered responses are the displacement and the stress, as shown in Figure 3-7 and Figure 3-8. The model calibration selected the observation data through the proposed experimental design method, as represented in Figure 3-12. To develop the sensitivity-based and proportionate bias calibration, the bound of unknown parameters E and P are [65, 75] and [1.5, 2.5]. With the bound information of unknown parameters, Table 4-3 determines the bound of model bias.

Table 4-3 The bound selection for sensitivity-based and proportionate bias

	E	W	Prediction	Sensitivity-based	Proportionate bias
σ	**LB (65)**	**LB (1.5)**	**2.1241e+07**	**9.5511e+04**	**0.9955**
	LB (65)	UB (2.5)	2.2065e+07	**9.2003e+05**	0.9583
	UB (75)	LB (1.5)	2.1241e+07	9.5464e+04	0.9955
	UB (75)	UB (2.5)	2.2065e+07	9.2000e+05	0.9583

	E	W	Prediction	Sensitivity-based	Proportionate bias
u	**LB (154.4)**	**LB (0.434)**	**0.049**	**-0.0106**	**0.784**
	LB (154.4)	UB (0.566)	0.038	0	1.000
	UB (245.6)	LB (0.434)	0.031	0.0074	1.239
	UB (245.6)	UB (0.566)	0.024	0.0144	**1.600**
	Lower Bound			**-0.14E+07**	**0.784**
	Upper Bound			**0.23E+07**	**1.600**

With the bound information of the unknown parameters and model bias, Table 4-4 gives the model calibration result. For model calibration, the calibration metrics in (4.7) to (4.9) are adopted. The result in Table 4-4 verifies that the proportionate bias calibration gives the lowest parameter estimation errors εrr_E^* and εrr_W^*.

Table 4-4 The model calibration result for automotive wheel rim model

	Without Bias	**Sensitivity-based**	**Proportionate bias**
δ	-	-0.0102	1.0175
E(Gpa)	180.5	245.6	**199.65**
err_E^*	0.0975	0.2280	**0.0017**
P (m)	0.4750	0.4750	**0.4833**
err_P^*	0.0500	0.0500	**0.0334**

4.3.3 Engineering Case Study 2: Automotive Steering Column Assembly FEM model

The purpose of this chapter is to demonstrate the efficacy of proportionate bias calibration method in a practical FEM model and experimental data. The difference of chapter 4.3.2 is that chapter 4.3.3 adopted real experimental data measured in experimental setup with practical measurement noise, and an industrial FEM model provided by Hyundai Motors group.

Automotive steering column is a device that helps a driver to change the driving direction of an automobile. One issue in the design of an automotive steering column is the desire to reduce the resonated vibration transferred from the engine or the roadway. Transferred vibration may make drivers uncomfortable. The design of an automotive steering column is based on the understanding of vibrational behaviors to avoid resonance. Therefore, the purpose of a computational model for an automotive steering column is to analyze the natural frequency and mode shape of the vibrations. In this chapter, we focus on natural frequencies matched to specific mode shapes that arise when a steering wheel vibrates in three axial bending directions. This approach is based on prior knowledge from industrial experts who designed the automotive steering column (offered to us via personal communication). These engineers have found that the vibrating strength of axial bending modes is the most powerful. The target modes of the natural frequency are the 1st, 2nd, and 4th modes. The 3rd mode of the natural frequency is not considered because it has a twisting mode shape.

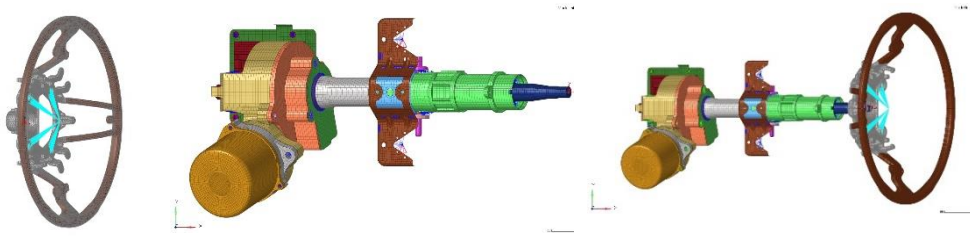
Figure 4-6 and Figure 4-7 show an automotive steering column and the computational model of an automotive steering column, respectively. An automotive

steering column consists of two sub-components: a steering wheel and a column assembly. Figure 4-6 (a) and Figure 4-7(a) show a steering wheel and Figure 4-6(b) and Figure 4-7(b) show a column assembly. The full steering column and its computational model are shown in Figure 4-6(c) and Figure 4-7(c), respectively. For computational modeling, Hypermesh 13.0 software is used as a pre-processor. 298,458 nodes and 214,268 elements are required for exact computation. To save computational costs, a computational model simplifies the geometric complexity of a real product. For example, the bolting system is simplified as a rigid body element (RBE2). The airbag in the middle of the steering wheel is simplified as a lumped mass with no geometric inputs. The geometry of the wheel cover uses shell elements, which only considers thicknesses as a geometric input variable.

Using the simplified steering column model, Autodesk Nastran 2018 is employed for a solver and post-processor. The automotive steering column model takes 575.66 seconds for the calculation of the natural frequencies and mode shapes. The complex design of an automotive steering column requires a significant amount of computational cost for analysis. Despite the calculation cost, however, the computational responses of the model are severely mismatched with experimental data. Therefore, the model requires optimization-based model calibration. For the optimization-based model calibration, the steering column model has five unknown parameters; elastic modulus of steel, density of a wheel cover, stiffness of an airbag, thickness of an airbag, thickness of ECU bracket, and thickness of a wheel cover. The bound information and initial value of each unknown parameters are as shown in

(a) (b) (c)

Figure 4-6 An automotive steering column: (a) Steering wheel, (b) Column assembly, (c) Entire steering column



(a) (b) (c)

Figure 4-7 An automotive steering column model: (a) Steering wheel, (b) Column assembly (c) Entire computational model of the steering column

Table 4-5 The initial value and bound information of unknown parameters in automotive steering column model

	Bound		Initial value
	Lower bound	Upper bound	
Density of the wheel cover (MPa)	0.50	1.50	1.00
Elastic modulus of column frame (MPa)	102500.00	307500.00	205000.00
Thickness of the wheel cover (mm)	2.75	8.25	5.50
Thickness of the ECU bracket (mm)	1.60	4.80	3.20

With the information about the unknown model parameters, the bound of sensitivity bias and proportionate bias can be determined by eq. (4.5) and (4.6). This chapter skips the whole bound selection process like Table 4-1 and Table 4-3, since the steering column model includes numerous unknown parameters to consider. The result of the bound is noted as Table 4-6.

Table 4-6 The bound information of bias term

Sensitivity-based bias		Proportionate bias	
Lower bound	Upper bound	Lower bound	Upper bound
-16.0524	18.4818	0.8340	1.6816

Similar to chapter 4.3.1 and 4.3.2, model calibration result specifies to OBMC with bound information, with sensitivity bias, and proportionate bias. The result of calibrated parameters are summarized in Table 4-7.

Table 4-7 The result of calibrated parameters

	OBMC	**OBMC_δ_{sb}**	**OBMC_δ_{pb} (Proposed)**
rho_wheel cover (MPa)	0.7614	0.8406	0.6391
T_wheel cover (mm)	5.5185	5.5256	5.5211
E_steel (GPa)	238.36	233.13	244.01
T_ECU Bracket (mm)	3.6224	3.6224	3.6220
K_airbag (N/mm)	1000.000	999.9999	999.9989
δ	-	1.0237	0.9749

Since all of practical study cannot figure out the real value of unknown parameters, the case study result in chapter 4.3.3 cannot show the parameter estimation errors

directly. Instead, the validity of the calibrated parameters is confirmed by the prediction errors of 2nd mode of natural frequency, whose observation data are not utilized for model calibration. (4.11) quantifies the prediction error of 2nd mode natural frequency.

$$\varepsilon = |\eta_{2nd\ M.F.} - \zeta_{2nd\ M.F.}|/\zeta_{2nd\ M.F.} \times 100\ (\%) \qquad (4.11)$$

With (4.1.1), the prediction errors by the calibrated parameters calculated in Table 4-7 are illustrated in Figure 4-8. The result implies that the calibrated result of proportionate bias can adjust the predicted responses to the observation data.

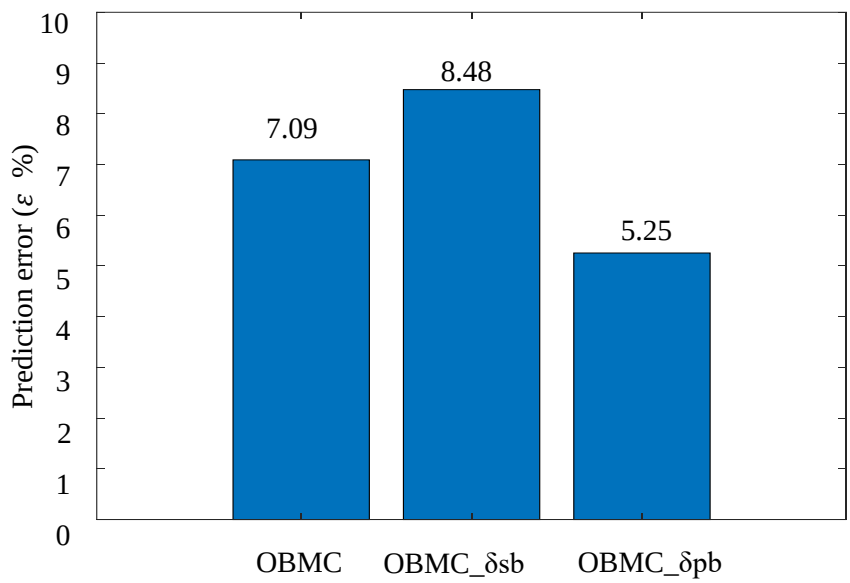


Figure 4-8 Prediction errors of 2nd mode natural frequency by the calibrated parameters

4.4 Summary and Discussion

The error sources in a computational and a physical model can degrade the parameter estimation errors in model calibration. Without considering error sources in model calibration, the estimates of unknown parameters might be unreasonable compared with physics-based information. To cope with the estimation errors of the unknown model parameters, Chapter 5 devices the proportionate bias calibration to consider the amount of uncontrollable errors in model calibration. The proportionate bias represents an adjusting ratio between observation and prediction due to the error sources. OBMC optimizes the proportionate bias term with unknown model parameters together to quantify the amount of proportionate bias. By the cantilever beam model and automotive wheel rim model for stress and displacement analysis, the accuracy of parameter estimation errors with proportionate bias calibration is compared with other model calibration method (e.g., model calibration without bias and sensitivity-based model calibration). The result shows the efficacy of the proportionate bias calibration.

The significant difference between proportionate bias calibration for OBMC and bias correction for Bayesian calibration is the required observation data. Bias correction aims to quantify the Gaussian process model for the discrepancy using the observation data measured from more than three design sites. The proportionate bias calibration method has a novelty that employs the experience-based bound information to support the insufficient observation data for quantifying the effect of error sources in model calibration. The proportionate bias prevents OBMC from optimizing the unknown parameters to the biased locations from the actual location to supplement the errors in predicted responses. Comparing with sensitivity-based

model calibration, the proportionate bias formulation is favorable for the model calibration by the multiple responses whose magnitude of each response value is severely different.

To expand use of the proposed method, it will be challenging to determine the validity of the bound information for unknown model parameters. In case studies, the author employs the statistical information of unknown parameters by the references, which has difficulty to gather the required information. In a practical engineering field, however, it may cause the degraded accuracy of the proposed method since the reliability of the bound information only depends on the accuracy of the industrial engineer's experience-based knowledge. Thus, there is a need to introduce the systematic method for considering epistemic uncertainties of the bound information as a future research.

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

2) **Hyejeong Son**, Guesuk Lee, Kyeonghwan Kang, Young-jin Kang, Byeng D. Youn, Ikjin Lee, and Yoojeong Noh, "Industrial Issues and Solutions to Statistical Model Improvement: A Case Study of an Automobile Steering Column," *Structural and Multidisciplinary Optimization*, Vol. 61, pp. 1739-1756, 2020.

Chapter 5

Comparison of Statistical Validation Metrics to Reduce Type II Errors in Model Validation

Statistical model validation (SMV) evaluates the prediction accuracy of computational models based upon observed data in a statistical manner (Hills and Trucano 1999) (Oberkampf and Barone 2006) (Oberkampf and Roy 2010) (Kat and Els 2012) (Sankararaman and Mahadevanb 2015) (Lee et al. 2019) (Kim and Youn 2019). SMV requires statistical validation metrics and hypothesis testing as tools for statistical decision-making of whether a computational model is valid or not. A statistical validation metric measures the discrepancy between predicted and observed results that is used for the decision-making process (Oberkampf et al. 2004) (Xiong et al. 2009) (Thonhofer et al. 2014) (Zhao et al. 2017). Hypothesis testing evaluates the plausibility of rejecting a null hypothesis according to a designated confidence level (Naylor and Finger 1967) (Balci and Sargent 1982) (Koch 2013) (Johnson et al. 2000) (Stanton et al. 2000) (Wilcox 2011) (Ross 2020). In SMV, the null hypothesis is that there is no significant discrepancy between the prediction and observation. Hypothesis testing constructs a distribution of the validation metric, assuming that the observation data belongs to the prediction's distribution. When the

value of the computed validation metric exists within the confidence level of this validation metric's distribution, the hypothesis testing is unable to reject the null hypothesis. Using this information, SMV determines the validity of the computational model.

Statistical validation metrics play a crucial role in SMV; they can change the result of the decision-making process. Under a variety of uncertainty sources, SMV can adopt a number of previously developed validation metrics in a statistical manner. Previous studies adopted the shape of the distribution function to consider a discrepancy in a global region (Kullback 1997) (Hills and Trucano 2002) (Mahadevan and Rebba 2005) (Oberkampf and Barone 2006) (Oberkampf and Barone 2006) (Ferson et al. 2008) (Jeon et al. 2015). The area metric is one popular metric. Developed by Ferson et al., the area metric generally employs a U-pooling method to solve practical problems using SMV (Ferson et al. 2008) (White and West 2019). Likelihood is a metric for goodness of fit, which measures the fidelity of a dataset to a designated distribution (Hills and Trucano 2002) (Oberkampf and Roy 2010) (Keysers et al. 2020). In other work, Bayes factor has been used to quantify the likelihood ratio of two possible distribution models for given data (Kass and Raftery 1995) (Berger and Mortera 1999) (Keysers et al. 2020). Kullback-Leibler divergence (KLD) has been used to imply the relative entropy of a probability density function with respect to another reference distribution (Kullback 1997) (Smith et al. 2006) (Pérez-Cruz 2008). Probability of separation (PoS) quantifies the separated degree of two distributions and is generally used for the classification of two datasets (Jeon et al. 2015). Probability residual is the squared form of the area discrepancy between two probability density functions (Oh et al. 2019) (Son et al.

2020).

To guarantee the accuracy of SMV, it is of importance to select a feasible statistical validation metric. Hypothesis testing must consider both Type I and Type II errors. Type I error denotes when an accurate prediction model is wrongly rejected. In contrast, Type II error describes the situation when an invalid prediction result is improperly accepted. The lack of observation data and distribution fitting errors are the main reasons for both types of errors. These errors severely occur when the variance of observation data is relatively more extensive than a prediction due to the statistical uncertainties in observation. In engineering practice, however, Type II errors in hypothesis testing can be much hazardous than Type I errors since an invalid prediction can be used in a variety of activities with potential safety concerns such as the design of automobiles and planes. Thus, SMV requires an assessment of statistical validation metrics to arrive at the options that give the less Type II errors in decision-making when the variance of observation and prediction have a severe discrepancy.

Scholars in related fields have discussed the features desired for statistical validation metrics, according to specific scenarios (Liu et al. 2011) (Ling and Mahadevan 2013) (Bi et al. 2017) (Maupin et al. 2018). These studies deal with the extended capability of statistical validation metrics for use with observation data that has particular characteristics. However, the validation metrics used in existing studies tend to be very sensitive to the distance between the observation and prediction and only weakly consider the discrepancy of the distribution with respect to the distributed degree of each dataset. For an accurate SMV, it is necessary for statistical validation metrics to quantify the difference between the distributions with

respect to not only the difference in mean, but also the variance. A discrepancy of variance between observation and prediction can occur in various situations. For example, when the uncertainty of the predicted response varies according to an input parameter, the distributions of the observed and predicted results will have a difference in variance.

To address this research need, this study focuses on statistical validation metrics, which enable capture of the variance difference between observation and prediction to reduce decision errors in SMV. In particular, the research outlined in this study considers invalid prediction models that arise due to unknown parameters, which represent the expected error sources in model prediction. The remainder of this study is organized as follows. Chapter 5.1 gives an overview of statistical validation metrics. Chapter 5.2 presents a detailed examination of the statistical validation metrics through a numerical case study that considers differences in uncertainty of the prediction and observation that arise due to unknown parameters. Discussion and demonstration of the results in Chapter 5.2 is in Chapter 5.3 why area metric and PoS was less sensitive when the variance of predicted responses are different. The engineering case study in Chapter 5.4 confirms the validity of the discussion in Chapter 5.3. The study concludes with a summary and remarks, provided in Chapter 5.5.

5.1 Brief Review of Statistical Validation Metrics

Among the developed statistical validation metrics, this study focuses on metrics that

use statistical distributions of the predictions and observations. Table 5-1 summarizes all equations of the considered statistical validation metrics, both S and L types. C_1 and C_2 denote the distributions of the data, which can be either observation or prediction. f and F, respectively, stand for a probability density function (PDF) and a cumulative density function (CDF) of C_1 or C_2. d denotes each data element, for the observation data. θ presents the parameters of the distributions, which are estimated from both observation and prediction. $\tilde{x}_{c_1}$ and $\tilde{x}_{c_2}$ in (5.7) are the medians. The term S in (5.8) is a scale parameter that is only used in the probability residual metric. The SMV considered in this study adopts hypothesis testing for decision making. This chapter provides comprehensive information about SMV that uses a hypothesis testing process. Furthermore, six statistical validation metrics are introduced in Chapter 5.1 to assess their capability for use with the hypothesis testing method.

Table 5-1 Comparison of the model calibration results of unknown variable E (Young's modulus)

Metrics	Type	Equation	
Area metric	S	$A(f_{C_1}, f_{C_2}) = \int_{-\infty}^{\infty} \lvert f_{C_1}(x) - f_{C_2}(x) \rvert dx$	(5.1)
		$A_u(F_1, F_2) = \int_{-\infty}^{\infty} \lvert F(u_{pred}) - F(u_{obs}) \rvert du$ where $u_i = F_{pred}(x_i)$	(5.2)
Likelihood	L	$Lh(\boldsymbol{\zeta}\lvert\boldsymbol{\theta}) = \prod_{i=1}^{n} f(\zeta_i\lvert\boldsymbol{\theta})$	(5.3)
		$\log Lh(\boldsymbol{\zeta}\lvert\boldsymbol{\theta}) = \ln(\prod_{i=1}^{n} f(\zeta_i\lvert\boldsymbol{\theta})) = \sum_{i=1}^{n} \ln(f(\zeta_i\lvert\boldsymbol{\theta}))$	(5.4)
Kullback-Liebler Divergence	S	$D_{KL}(c_1\lvert c_2) = \int_{-\infty}^{\infty} f_{c_1}(x) \ln\left(\frac{f_{C_1}(x)}{f_{C_2}(x)}\right) dx$	(5.5)
Bayes factor	L	$B_0 = \frac{p\{\boldsymbol{\zeta}\lvert H_o: \theta_{pred} = \theta_{obs}\}}{p\{\boldsymbol{\zeta}\lvert H_1: \theta_{pred} \neq \theta_{obs}\}} = \frac{Lh(\boldsymbol{\zeta}\lvert\theta_{obs})}{\int\int Lh(\boldsymbol{\zeta}\lvert\theta) f^{prr}(\theta) d\mu d\sigma}$	(5.6)
Probability of separation	S	$PoS = \frac{(e^{1-2P_{NS}} - 1)}{e - 1}$ where, $P_{NS} = \int_{-\infty}^{\infty} F_{c_2}(x) f_{c_1}(x) dx$ for $\tilde{x}_{c_1} \leq \tilde{x}_{c_2}$	(5.7)
Probability residual	S	$PR = S \times \int_{-\infty}^{\infty} \left(f_{C_1}(x) - f_{C_2}(x)\right)^2 dx$	(5.8)

5.1.1 Area metric

The area metric is an integral of the difference between two CDFs, as stated in (5.1). Ferson et al. recommended using the area metric with a U-pooling method (Ferson et al. 2009). The research described in this study adopted the area metric with a U-pooling method for a comparison study. The area metric with U-pooling is as shown in (5.2). The advantage of the U-pooling method is that it can integrate the data achieved in a variety of environments for validation (Ferson et al. 2009) (Liu et al. 2011). Furthermore, the metric is enabled when only a few observation data are available; in fact, even if only one data sample exists. The area metric with U-pooling varies from 0 to 0.5.

5.1.2 Likelihood

Likelihood in (5.3) measures the multiplication of all prediction PDFs at the observation sites (Edwards 1984) (Severini 2000) (Myung 2003). For practical reasons, this study applies a logarithmic transformation to the likelihood, as shown in (5.4). The multiplication of PDF values lower than one makes the likelihood smaller, which is not beneficial to the computation. The likelihood has no limits on its ability to increase or decrease, which helps to maximize the likelihood for parameter estimation. Since the likelihood metric uses observation data without a fitting method for the distribution, it gives a relatively good validation accuracy with even a small number of observation data. Likelihood can estimate differences that arise due to the variance of the observed and predicted results.

5.1.3 Kullback-Leibler Divergence (KLD)

Kullback-Leibler divergence (KLD), as presented in (5.5), is motivated by information systems that quantify the entropy of given data (Kullback and Leibler 1951) (Kullback 1997) (Anderson and Burnham 2004) (Bishop 2006). This aspect makes KLD sensitive to the variance difference between the two datasets. KLD is asymmetric; thus, if the distributions of fc_1 and fc_2 are inverted, different values are calculated. Therefore, for the case study described in Chapter 5.2, fc_1 is substituted into the observation distributions and fc_2 is used for the prediction distributions. The metric becomes zero if and only if the distributions of C_1 and C_2 are the same in a global domain at x. In contrast, a high KLD value indicates the separation of the two distributions.

5.1.4 Bayes Factor

The Bayes factor forms a ratio of the null and alternative hypothesis distributions (Berger and Mortera 1999) (Morey and Rouder 2011) (Ling and Mahadevan 2013) (Keysers et al. 2020). For the probabilities of the null hypothesis and the alternative hypothesis, the Bayes factor in (5.6). adopts the likelihood of observation data under the prediction model parameters' prior distribution (Liu et al. 2011). The research described in this study assumes only two model parameter values estimated from the observation and prediction data are possible, since the numerical example has no specific prior information. Thus, $f^{prr}(\theta)$ becomes 0.5. This assumption affects the accuracy of the SMV. However, I want to note that some practical problems have no prior information to enable use of the Bayes factor.

5.1.5 Probability of Separation (PoS)

The probability of separation (PoS) metric originates from the probability of failure, which is generally used in reliability analysis (Jeon et al. 2015). The PoS assumes that the median of C_1 must be smaller than the median of C_2. This is because the probability of failure is defined when the overall strength of a system is smaller than the load applied to the system. From the derivation in (5.9), the P_{NS} is defined to be within the range of [0, 0.5]. *PoS* is the normalization of P_{NS}, which is defined in the range of [0, 1]. The (5.9) explains that the maximum value of P_{NS} becomes 0.5 when the distribution of C_1 and C_2 is the same.

$$P_{NS|c_1=c_2} = \int_{-\infty}^{\infty} F_{c1}(x) f_{c1}(x) dx = F_{c1}(x)F_{c1}(x)|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} f_{c1}(x)F_{c1}(x)dx \quad (5.9)$$

$$\int_{-\infty}^{\infty} F_{c1}(x) f_{c1}(x) dx = \frac{1}{2} F_{c1}(x)^2|_{-\infty}^{\infty} = \frac{1}{2}(1-0) = 0.5$$

5.1.6 Probability Residual (PR)

The probability residual (PR), shown in (5.8), is a discrepancy measure for statistical model calibration that seeks to reinforce the limits of low sensitivity observed in existing discrepancy measures (Lee et al. 2018) (Oh et al. 2019). Although the PR has not been used as a discrepancy measure for SMV, this study examines the possibility of using it for this purpose. The minimal value of PR is zero, which occurs when the prediction and observation are perfectly in the same distribution. PR is not limited in its ability to increase the metric value. PR uses the distribution function that is estimated by a distribution fitting method. Thus, distribution fitting error can

occur, especially when the number of observations is small. The validation result described in Chapter 5.3 supports the idea that PR shows low accuracy when the number of observations is less than 10.

5.2 A comparison study of statistical validation metrics

Chapter 5.2 discusses the comparison of statistical validation metrics, with discrepancies in variance across the observed and predicted results. As previously discussed, the existing studies concentrate on examining the SMV accuracy when the mean of prediction and observation severely occurs. Since the variance discrepancy also causes hypothesis testing errors when it is not validated correctly, the statistical validation metrics require to be evaluated quantify the discrepancy of observation and prediction properly.

5.2.1 Problem definition

A nonlinear response with statistical input parameters is defined for the comparison study, as shown in (5.10) (Youn et al. 2008)

$$G(X,\theta) = e^{\left(-\frac{1}{1+100X^2+2\theta^2+X^2\theta^2}\right)} + \varepsilon \qquad (5.10)$$

where θ denotes the unknown variable that causes the discrepancy between the observation and prediction. The X and ε are the model input parameters. The nonlinearity in function G(X, θ) creates a change of variance as well as mean values

according to the statistical variables X and θ. Through the function G(X, θ) in (5.10), this study considers the importance of the change in variance for validating a nonlinear response that numerous engineering examples have. Furthermore, the case studies reflect Type II errors by assuming that the prediction and observation have different θ values. For a prediction model, Type II errors are much more dangerous because an invalid prediction of this type might be utilized in a practical field with catastrophic results.

This study defines two case studies that were explored to compare statistical validation metrics. Case 1 deals with the situation in which the variance of the observed and predicted results is the same. Case 2 considers different variances in the observation and prediction. Case 2 is further broken down into Case 2-1 and 2-2. Case 2-1 examines the situation in which the variance of the observation is smaller than that of the prediction. Case 2-2 explores the situation where the variance of the observation is larger than that of the prediction. Table 5-2 summarizes the input parameters X, θ, and ε in each case study. Case 1 specifies that ε follows a Gaussian distribution with zero-mean; variance is specified as 0.08, 0.12, 0.16, and 0.2. Case 2 considers the statistical input parameter X, instead of the measurement error ε. X follows a Gaussian distribution; the mean is 0.1 and variance is specified as 0.08, 0.12, 0.16, and 0.2.

Depending on the value of θ, the mean and variance of the prediction and observation can be different. Figure 5-1 depicts the graphs of the function G(X, θ), concerning the unknown variable θ. The solid black line is the mean of the function G(X, θ). The dashed line, dotted line, dash-dotted line, and solid line with a star mark show the four different uncertain ranges, which give the 10 and 90% CDF values.

Figure 5-1 (a) is the graph of Case 1, which has the same variance, regardless of θ. Figure 5-1 (b) is the graph of the function G(X, θ) in Case 2, which has a different variance about θ. The red vertical line denotes the predicted response at θ =1. The blue vertical line at θ =0 shows the observation for Case 2-1, and the line at θ =2 shows the observation for Case 2-2. The graphs in Figure 5-1 show that the uncertain range of all the function G(X, θ) increases when the variance of ε and X increases.

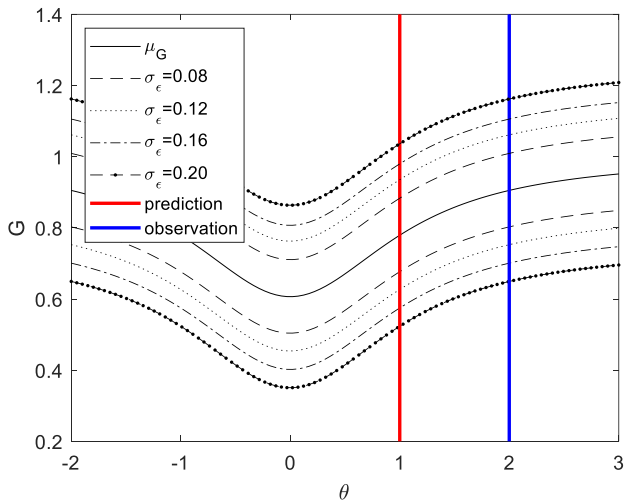
The observation dataset includes 3 to 30 sampled data to explore the effect of the number of observations. 10^4 datasets for 3 to 30 sampled data are prepared to repeat the test. The prediction includes 10^5 sampled data to represent the distribution of prediction without statistical uncertainty. Since the prediction and observation data were sampled from the different values of θ, the SMV result should be that the prediction model is invalid. Therefore, the accuracy of the validation metrics denotes the number of observation datasets for which the validation metric value is outside of an acceptable range. The accuracy of the validation metrics is defined in (5.11).

$$The\ \ rejection\ \ ratio = \frac{n}{N} \tag{5.11}$$

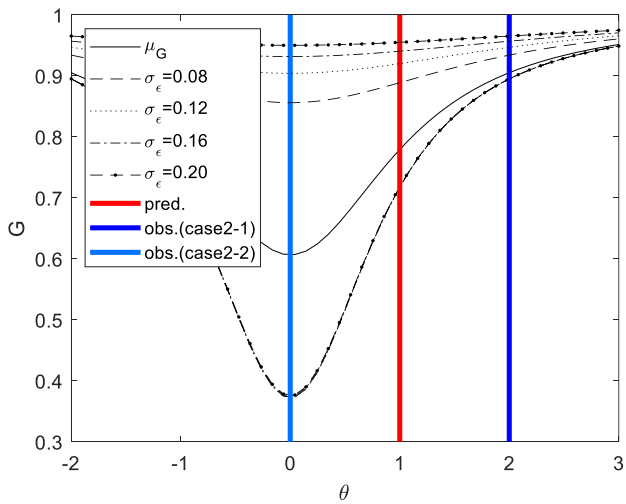
N is the number of datasets, 10^4, and n is the number of datasets located outside of the acceptable range. The validation metric, which gives a high value for the rejection ratio, shows a favorable result in the SMV.

Table 5-2 The model parameters and corresponding true values for the cantilever beam

		X	θ	ε	$S.t.d$
Case 1	Prediction	0.1	1	N~(0, 0.08 to 0.2^2)	$S.t.d_{\text{Pred.}} = S.t.d_{\text{obs}}$
	Observation	0.1	2	N~(0, 0.08 to 0.2^2)	
Case 2-1	Prediction	N~(0.1, 0.08 to 0.2^2)	1	-	$S.t.d_{\text{Pred.}} > S.t.d_{\text{obs}}$
	Observation	N~(0.1, 0.08 to 0.2^2)	2	-	
Case 2-2	Prediction	N~(0.1, 0.08 to 0.2^2)	1	-	$S.t.d_{\text{Pred.}} < S.t.d_{\text{obs}}$
	Observation	N~(0.1, 0.08 to 0.2^2)	0	-	



(a)

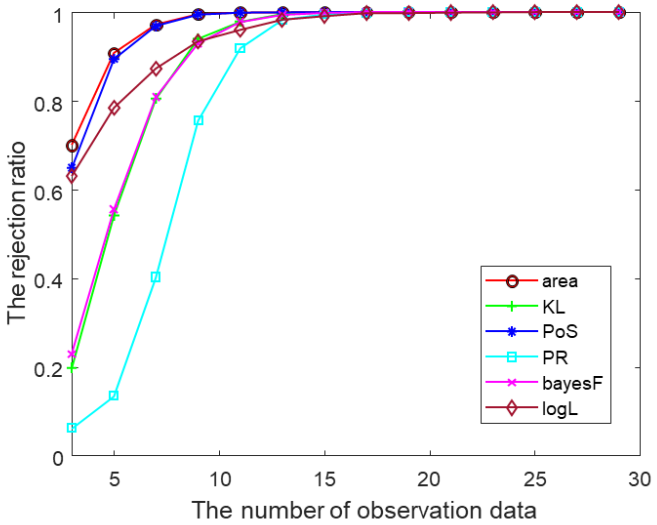


(b)

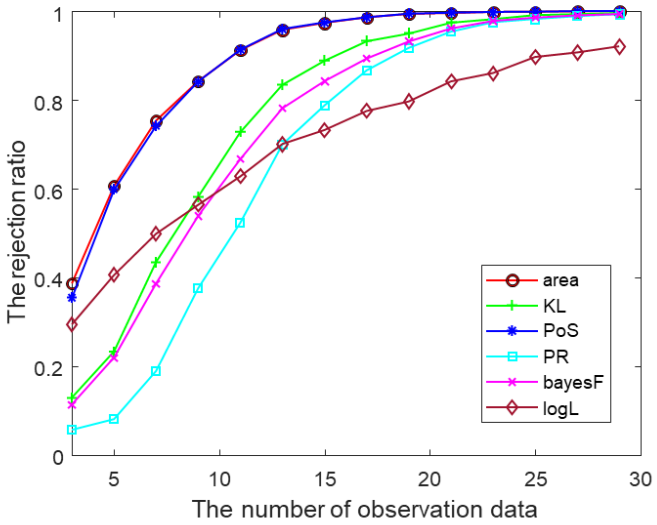
Figure 5-1 Predicted and observed responses (a) Case 1 (b) Case 2

5.2.2 Results of statistical model validation accuracy

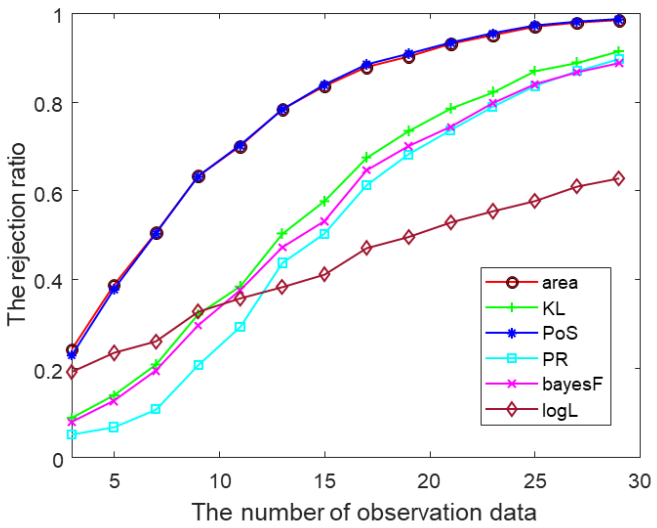
Figure 5-2, Figure 5-3, and Figure 5-4 show the rejection ratio based on the number of observations. All figures provide that the rejection ratio increases as the number of observations increases. As the uncertainty of input parameter ε or X increases, the rejection ratio requires more observation data for the rejection ratio to reach one. A large amount of uncertainty in the function G(X, θ) results in a lower rejection ratio, since the overlapped area of the prediction and observation distributions prevents distinguishing between the two distributions' discrepancy. The following paragraphs explain the results of Case 1, Case 2-1, and Case 2-2, as shown in Figure 5-2, Figure 5-3, and Figure 5-4.


The rejection ratio
The number of observation data
area
KL
PoS
PR
bayesF
logL

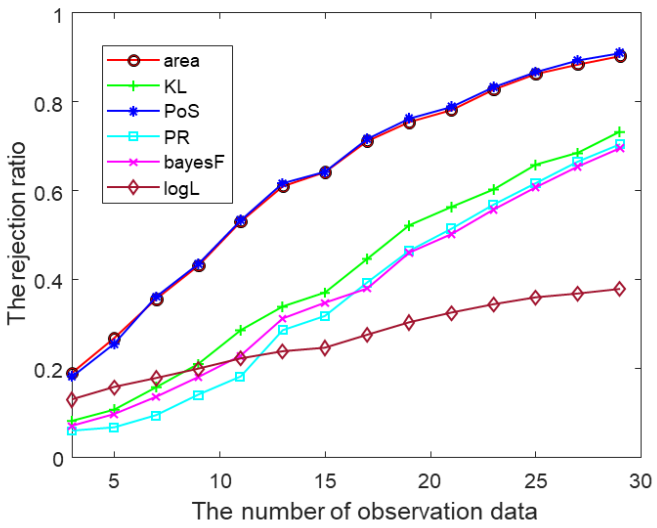
(a)


The rejection ratio
The number of observation data
area
KL
PoS
PR
bayesF
logL

(b)

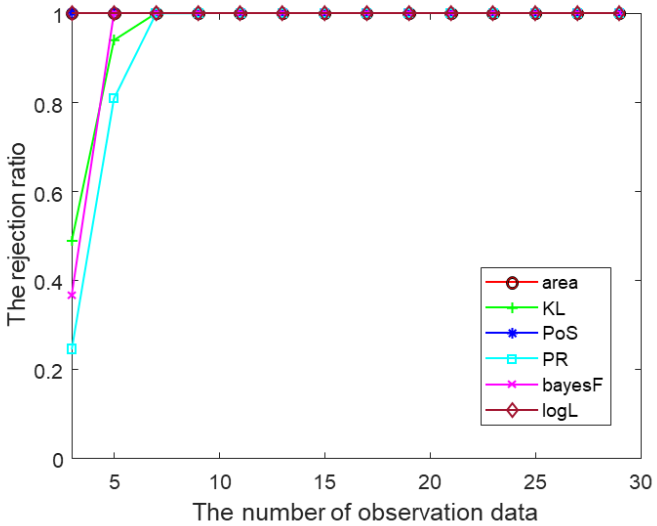


(c)

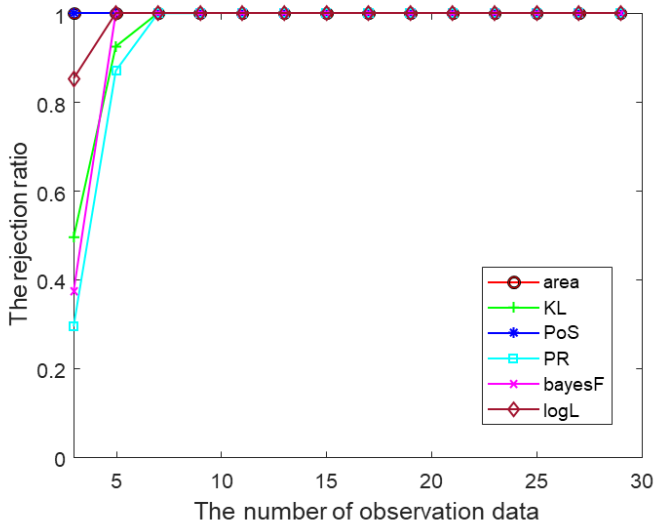


(d)

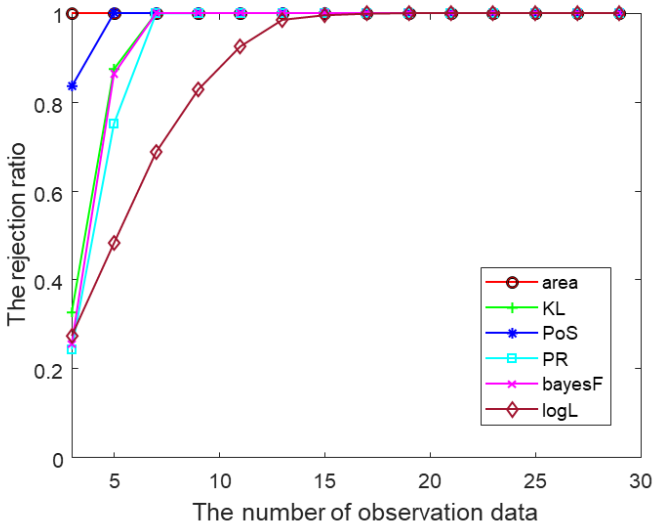
Figure 5-2 Case 1: The rejection ratio for the number of observation data under different uncertainties of ε (a) ε~N(0, 0.082), (b) ε~N(0, 0.122), (c) ε~N(0, 0.162), and (d) ε~N(0, 0.202)


The rejection ratio
The number of observation data
area
KL
PoS
PR
bayesF
logL

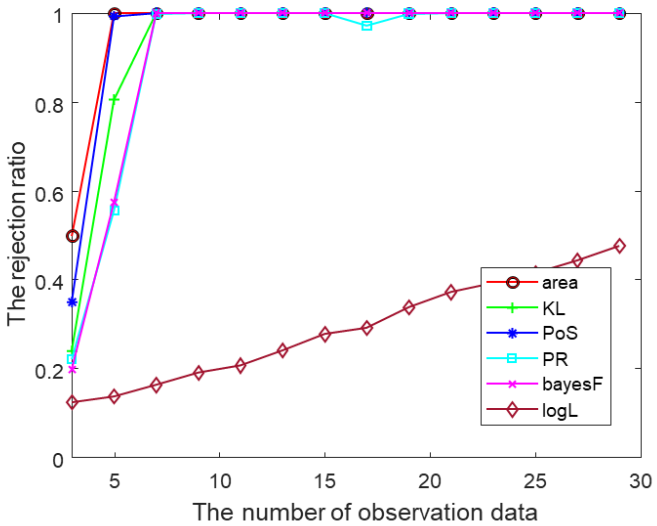
(a)


The rejection ratio
The number of observation data
area
KL
PoS
PR
bayesF
logL

(b)

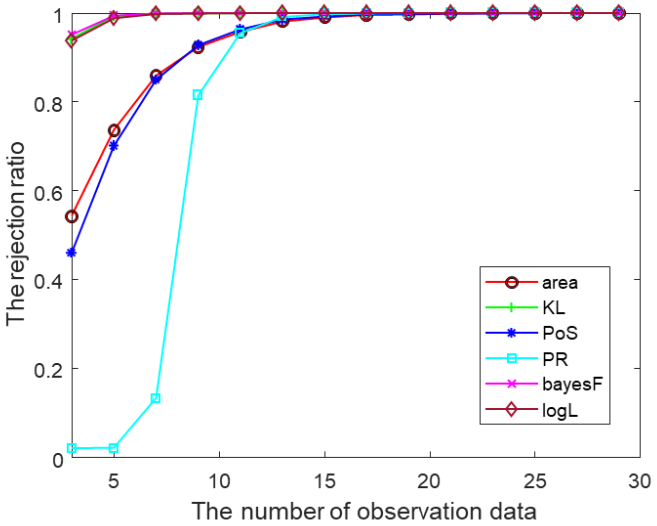


(c)

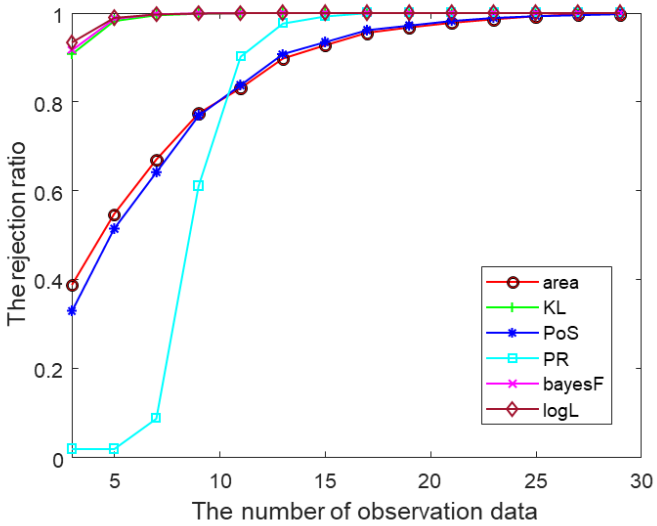


(d)

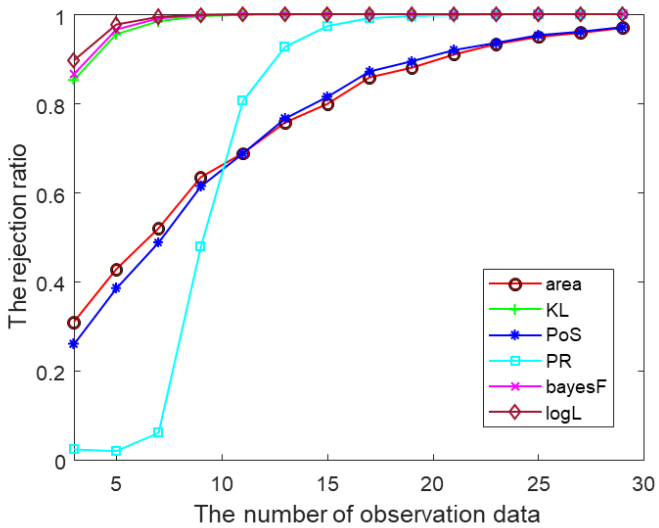
Figure 5-3 Case 2-1: The rejection ratio for the number of observation data sampled at the site θ=2 under different uncertainties of X (a) $X \sim N(0.1, 0.08^2)$, (b) $X \sim N(0.1, 0.12^2)$, (c) $X \sim N(0.1, 0.16^2)$, and (d) $X \sim N(0.1, 0.20^2)$


The rejection ratio
The number of observation data
area
KL
PoS
PR
bayesF
logL

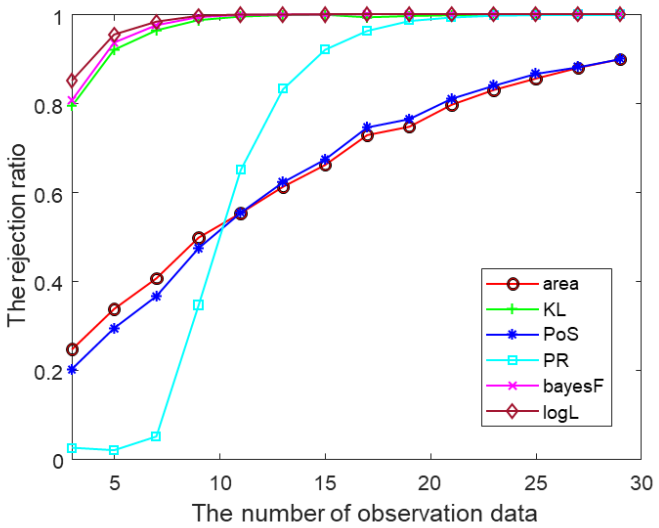
(a)


The rejection ratio
The number of observation data
area
KL
PoS
PR
bayesF
logL

(b)



(c)



(d)

Figure 5-4 Case 2-2: The rejection ratio for the number of observation data sampled at the site θ=0 under different uncertainties of X (a) X ~N(0.1, 0.08^2), (b) X ~N(0.1, 0.12^2), (c) X ~N(0.1, 0.16^2), and (d) X ~N(0.1, 0.20^2)

i. Case 1: $S.t.d_{\text{Pred.}} = S.t.d_{\text{obs.}}$

According to the graphs shown in Figure 5-2, the rank of the rejection ratio for the studied validation metrics is determined as follows: area metric≈PoS > KLD > Bayes factor > PR≈likelihood. Overall, the area metric and PoS give the highest rejection ratio. The validation metric with the second highest rejection ratio is KLD. Although the rejection ratio of KLD is almost the same as the Bayes factor's, the rejection ratio is lower than KLD's when the uncertainty of the function G(X, θ) increases. PR gives the lowest rejection ratio with a small number of observations (i.e., under 10). In contrast, likelihood shows the lowest rejection ratio with a large overall number of observation data (i.e., over 10).

ii. Case 2: $S.t.d_{\text{Pred.}} \neq S.t.d_{\text{obs.}}$

As explained in Chapter 5.2.1, Case 2 is divided into Cases 2-1 and 2-2. In Case 2-1, the rejection ratio of all statistical validation metrics, except for the likelihood, reached one before the number of observation data exceeded 10, as shown in Figure 5-3. This is because a small variance of observed responses reduces the statistical uncertainties for the observation results. In Case 2-2, the rejection ratio in Figure 5-4 indicates that the KLD, Bayes factor, and likelihood show a high rejection ratio. PR shows the lowest rejection ratio with a small number of observation data. Overall, the KLD and Bayes factor give the highest rejection ratio, as shown in Figure 5-3 and Figure 5-4, except for the case of a small number of observations. When only a small observation (i.e., less than five) is available, the area metric and PoS give a higher rejection ratio than KLD and Bayes factor.

Under unknown parameters in model prediction, it is common for an engineering system response to have different amounts of variance. Thus, the comparison result can suggest that KLD or Bayes factor can be a useful discrepancy measure for the variance discrepancy. Since KLD is a metric to quantify the relative entropy between two distributions, it is advantageous to distinguish the difference in the distributions. The area metric and PoS are recommended when a limited number of observations is available. If the number of observations is small, the statistical uncertainty in observations can cause a discrepancy in the variance between observation and prediction. Therefore, when the number of observations is small, it is desirable to use a metric that is less sensitive to the differences of variance. One noticeable point is that the Bayes factor's accuracy is related to the accuracy of the prior information. The Bayes factor is recommended when reliable prior information is available.

5.3 Discussion and Demonstration

This chapter explains why the area metric and PoS have low accuracy for variance discrepancy by deriving a possible metric range. Chapter 5.3.1 graphically explains the area metric's possible range when the mean and variance of the two distributions are increased. Chapter 5.3.2 discusses the possible range of KLD by examining the limit calculation when the mean and variance of the two distributions increase.

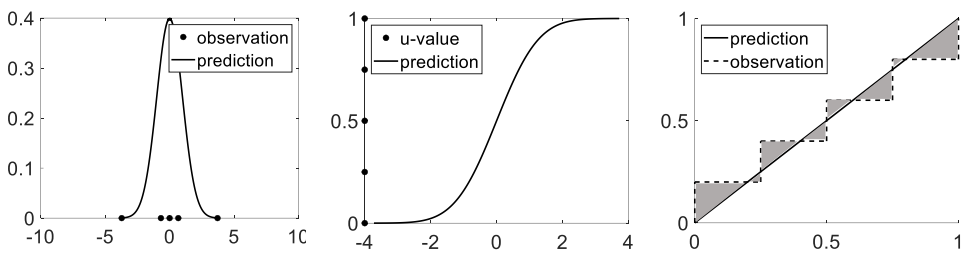
5.3.1 Discussion about the low accuracy of the area metric in a variance

change

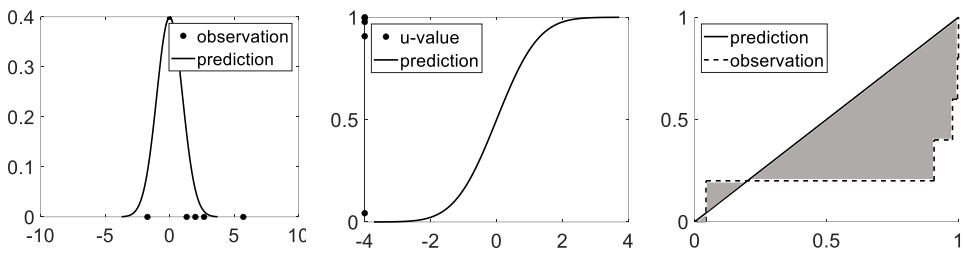
Figure 5-5 and Figure 5-6 illustrate the process for calculation of the area metric when the distribution of observation and prediction have a discrepancy in the mean and variance. The left figure is the prediction PDF and the observation data. The figure located in the middle of Figure 5-5 and Figure 5-6 shows the prediction CDF and the u-value of the observation. The graph on the right side is the empirical CDF of the u-value and the prediction CDF values. illustrates the change in the area metric when the mean of the observation is increased. The observation and prediction in Figure 5-5 (a) follow a standard normal distribution and the mean of the observation is increased from three to six in Figure 5-5 (b) and (c). The calculated area metric is increased from 0.07 to 0.5, according to the increased discrepancy of the mean. When the distributions of the prediction and observation are the same, the u-value of the observation is uniformly distributed, as shown in the second graph in Figure 5-5 (a). As the discrepancy of the mean between the distribution of the observation and prediction increases, the u-values of the observation move to the right side, as shown in Figure 5-5 (b) and (c). When all of the u-values shift to one, the area discrepancy of the empirical CDF between the prediction and observation reaches 0.5. In contrast, the u-value of the observation shifts to zero as the mean of the observation data decreases. In that case, the final area metric becomes zero. Overall, the discrepancy of the mean shifts all u-values to zero or one. This result in the area metric values can be in a range of [0, 0.5], the area metric's theoretical minimum and maximum.

On the other hand, the range of the metric is reduced in the variance difference. Figure 5-6 shows the area metric calculation process with U-pooling in which the variance of observation is increased. The original observation and prediction in

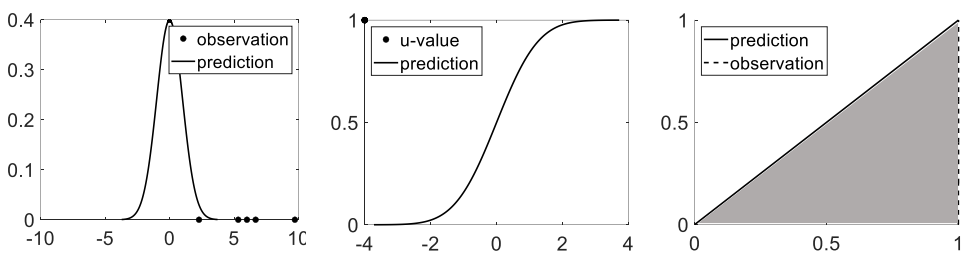
Figure 5-6 (a) follow a standard normal distribution. The observation variance in Figure 5-6 (b) and (c) is increased from one to three and six. Unlike Figure 5-5, the u-values depicted in the second graph of Figure 5-6 partially converge to a different location, such as 0, 0.5, and 1. The area metric's possible range cannot reach the minimum or maximum because u-values cannot converge to one location, 0 or 1. Thus, the area metric of the observation in Figure 5-6 (c) is calculated as 0.17. Under the discrepancy of variance, the range of the area metric depends on the number of observations. In conclusion, this research explains that the area metric with U-pooling falls within a short range of metric values under the variance discrepancy. Due to the large variance of the observation, the area metric value became smaller and within the valid range, even though observation datasets follow invalid conditions.



(a)

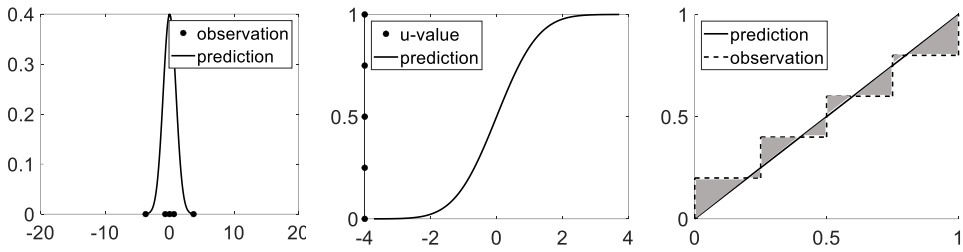


(b)



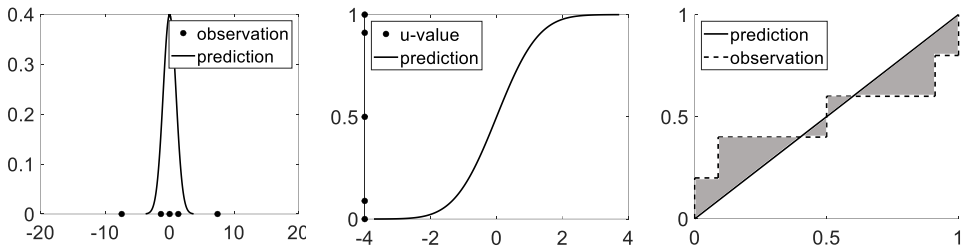
(c)

Figure 5-5 The calculation of the area metric when the difference of mean is increased (a) Observation N~(0,1^2) (b) Observation N~(2,1^2) (c) Observation N~(6,1^2) (Prediction follows a standard normal distribution)



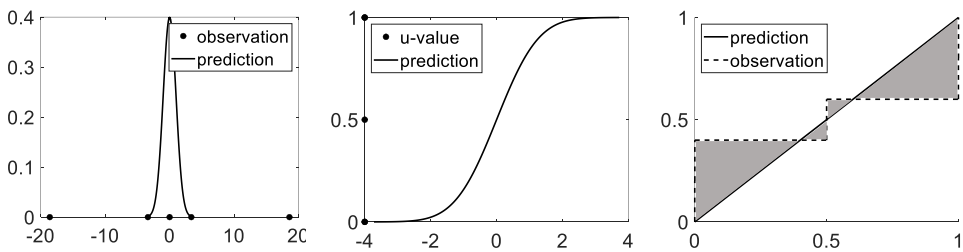
Au = 0.07

(a)



Au =0.13

(b)



Au =0.17

(c)

Figure 5-6 The calculation of the area metric when the difference of variance is increased (a) Observation N~(0,1^2) (b) Observation N~(0,2^2) (c) Observation N~(0,6^2) (Prediction follows a standard normal distribution)

5.3.2 Discussion about the low accuracy of the Probability of Separation (PoS) in a variance change

This chapter focuses on obtaining the limit of PoS as the mean or variance discrepancy approaches infinity. The predicted and observed responses are assumed to follow a normal or lognormal distribution. This assumption can consider both symmetric and asymmetric distributions. Appendix A provide a detailed derivation of P_{NS} in the PoS equation to derive the analytical metric's limit. (A.7) is the P_{NS} when the predicted and observed responses follow a normal distribution. μ_1 and μ_2 denotes the mean value. σ_1 and σ_2 stands for S.t.d. The mean discrepancy approaches infinity, which is the same as when μ_2 approaches infinity or μ_1 approaches minus infinity. The limit of P_{NS} as μ_1 approaches minus infinity is (5.12). The limit of P_{NS} as μ_1 approaches minus infinity is shown in (5.13).

$$\lim_{\mu_1 \to -\infty} P_{NS} = \lim_{\mu_1 \to -\infty} \Phi\left(\frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right) = \Phi\left(\lim_{\mu_1 \to -\infty} \frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right) = \Phi(-\infty) \quad (5.12)$$
$$= 0$$

$$\lim_{\mu_2 \to \infty} P_{NS} = \lim_{\mu_2 \to \infty} \Phi\left(\frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right) = \Phi\left(\lim_{\mu_2 \to \infty} \frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right) = \Phi(-\infty) = 0 \quad (5.13)$$

As the discrepancy of the mean increases, the P_{NS} converges to zero. Thus, the limit of PoS becomes one, as shown in (5.14).

$$\lim_{\substack{\mu_1\to-\infty \text{ or}\\ \mu_2\to\infty}} \text{PoS} = \lim_{\substack{\mu_1\to-\infty \text{ or}\\ \mu_2\to\infty}} \frac{e^{(1-2P_{NS})}-1}{e-1}$$
$$= \frac{1}{e-1}\left[\exp\left(1-2\times\left(\lim_{\substack{\mu_1\to-\infty \text{ or}\\ \mu_2\to\infty}} \text{P}_{NS}\right)\right)-1\right] \qquad (5.14)$$
$$= \frac{e^{(1-0)}-1}{e-1} = 1$$

The result in (5.14) shows that PoS can reach its maximum when the discrepancy of the mean in the observation and prediction increases. However, when the discrepancy of the variance increases, the possible PoS range is different. (5.15) offers the limit of P_{NS} as the variance σ_1 or σ_2 approach infinity. The limit of P_{NS} as the variance approaches minus infinity is an unreal situation because the variance is always a positive number.

$$\lim_{\sigma_1 \text{or} \sigma_2\to\infty} P_{NS} = \lim_{\sigma_1 \text{or} \sigma_2\to\infty} \Phi\left(\frac{\mu_1-\mu_2}{\sqrt{\sigma_1^2+\sigma_2^2}}\right) = \Phi\left(\lim_{\sigma_1 \text{or} \sigma_2\to\infty} \frac{\mu_1-\mu_2}{\sqrt{\sigma_1^2+\sigma_2^2}}\right) = \Phi(0) \qquad (5.15)$$
$$= 0.5$$

Using (5.15), the limit of P_{NS} converges to 0.5, which cannot be zero. Thus, the hypothesis testing accepts some invalid observation datasets with a large discrepancy of variance that is due to the high P_{NS}.

In the case of the log-normality conditions, (A.17) gives the analytical equation of P_{NS}. To consider the situation that the μ_2 and σ_2 approach infinity, Appendix 2 involves transforming the lognormal distribution's model parameters to the μ_2 and σ_2 using (A.15) and (A.16). (5.16) gives the limit of P_{NS} as μ_2 approaches infinity.

$$\lim_{\mu_2\to\infty} P_{NS} = e^{\frac{s_1^2}{2}+m_1}\Phi\left(\lim_{\mu_2\to\infty}\frac{m_1-\ln(\frac{\mu_2^2}{\sqrt{\sigma_2^2+\mu_2^2}})-s_1^2}{\sqrt{s_1^2+\ln(\frac{\sigma_2^2}{\mu_2^2}+1)}}\right) = e^{\frac{s_1^2}{2}+m_1}\Phi(-\infty) \qquad (5.16)$$
$$= 0$$

As with the result of (5.14) and (5.15), (5.16) shows that the limit of P_{NS} reaches zero, and PoS can reach its maximum when the discrepancy of the mean in the observation and prediction increases. In contrast, the limit of P_{NS} as σ_2 approaches infinity is non-zero. This is calculated by L'Hospital's rule, as shown in (5.17).

$$\lim_{\sigma_2\to\infty} P_{NS} = e^{\frac{s_1^2}{2}+m_1}\Phi\left(\lim_{\sigma_2\to\infty}\frac{m_1-\ln\left(\frac{\mu_2^2}{\sqrt{\sigma_2^2+\mu_2^2}}\right)-s_1^2}{\sqrt{s_1^2+\ln\left(\frac{\sigma_2^2}{\mu_2^2}+1\right)}}\right)$$
$$= e^{\frac{s_1^2}{2}+m_1}\Phi\left(\lim_{\sigma_2\to\infty}\frac{\frac{\sqrt{\sigma_2^2+\mu_2^2}}{\mu_2^2}\times\mu_2^2\times\frac{1}{2}(\sigma_2^2+\mu_2^2)^{-\frac{3}{2}}\times 2\sigma_2}{\frac{1}{2}\left(s_1^2+\ln\left(\frac{\sigma_2^2}{\mu_2^2}+1\right)\right)^{-\frac{1}{2}}\times\frac{\mu_2^2}{\sigma_2^2+\mu_2^2}\times\frac{2\sigma_2}{\mu_2^2}}\right) \qquad (5.17)$$
$$= e^{\frac{s_1^2}{2}+m_1}\Phi\left(\lim_{\sigma_2\to\infty}\left(s_1^2+\ln\left(\frac{\sigma_2^2}{\mu_2^2}+1\right)\right)^{\frac{1}{2}}\right) = e^{\frac{s_1^2}{2}+m_1}\Phi(\infty) = e^{\frac{s_1^2}{2}+m_1}$$

The term $e^{\frac{s_1^2}{2}+m_1}$ is transformed to μ_1 using (A.15) and (A.16).

$$\lim_{\sigma_2 \to \infty} P_{NS} = e^{\frac{s_1^2}{2}+m_1} = e^{0.5\, ln\left(\frac{\sigma_1^2}{\mu_1^2}+1\right)+ln\left(\frac{\mu_1^2}{\sqrt{\sigma_1^2+\mu_1^2}}\right)} = \mu_1 \tag{5.18}$$

Using (5.18), the limit of PoS is as follows.

$$\lim_{\sigma_2 \to \infty} PoS = \frac{e^{(1-2\mu_1)}-1}{e-1} \tag{5.19}$$

The result in (5.19) gives that the limit of PoS can reach the value upper than one, which means PoS can give its maximum when the variance discrepancy increases, only when $\mu_1 < 0$. However, the negative mean value is impossible for the lognormal distribution. Therefore, the PoS metric can give a validation error in the hypothesis test because the metric's possible range decreases when the variance discrepancy between the observation and prediction increases.

5.4 Case Study

To demonstrate the result in chapters 5.3.1 and 5.3.2, chapter 5.4 employs an automotive wheel rim FEM model as introduced in chapter 3.3.2. The automotive wheel rim is a frame of a wheel that combines the wheel to the body of an automobile. This chapter considers a structural analysis to present maximum stress when a wheel rim is supported by the weight of an automobile and the pressure in a tire. Figure 3-7(a) illustrates the result of the stress analysis that the maximum stress (σ) is

colored in yellow. Figure 3-9 indicates the boundary and loading condition.

Table 5-3 summarizes the related parameters in this model. Among all model parameters, the load on the wheel is a statistical parameter that affects the uncertainties in maximum stress. By assuming that the statistical parameter follows normal or lognormal distribution, the value of maximum stress can be either symmetric or asymmetric distribution.

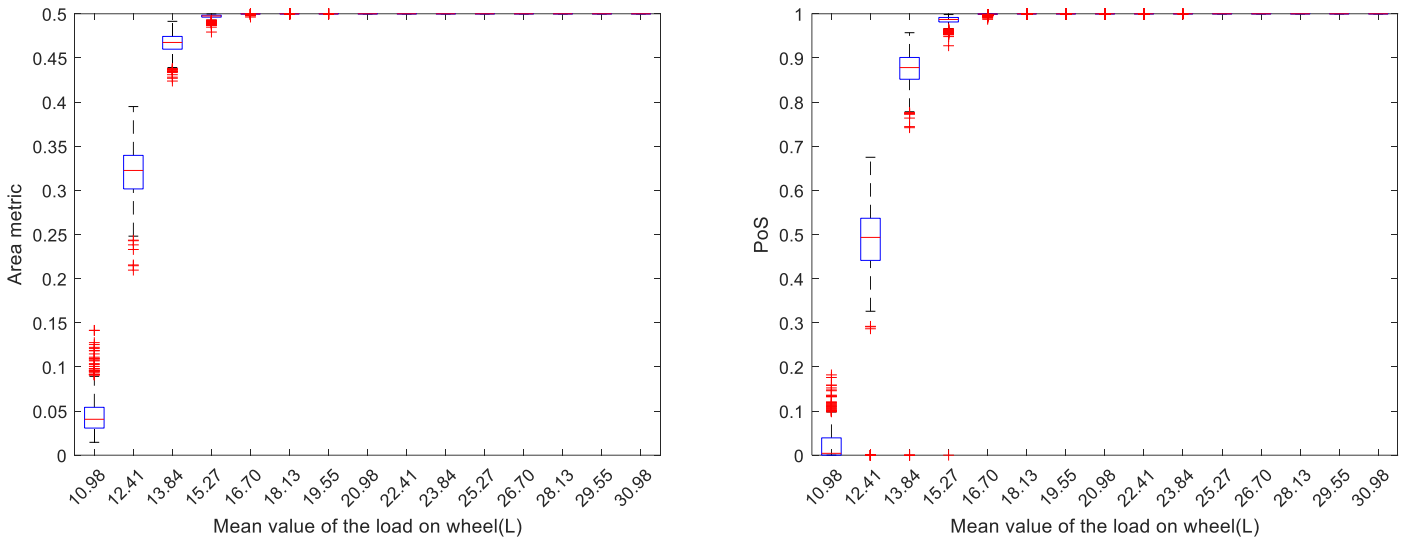
Parameter	Symbol (Unit)	Type	Mean	Standard deviation
Inflation pressure	P (bar)	Deterministic	2	-
Load on a wheel	L (kN)	Statistic	10.98	1.098
Young's modulus	E (GPa)	Deterministic	70	-
Density	P (kg/m^3)	Deterministic	2700	-
Poisson's ratio	v (.)	Deterministic	0.33	-

Table 5-3 The statistical information of input parameters in an automotive wheel rim FEM model

To show the increase of area metric and PoS when the prediction is inconsistent with observation, the mean and standard of deviation (S.t.d.) of the load on the wheel for observation increased until the area metric, and PoS converges. The mean increased up to 30.98kN, and the S.t.d. increased up to 52.10kN. For the distribution of observation and prediction, the number of observation and prediction data are 50 and 104. The observation has 1000 sets to repeat the metric calculation. With the 1000 metric values, Figure 5-7 to Figure 5-10 give the boxplot. Figure 5-7 and Figure 5-8 show the limit of area metric and PoS when the mean of L in the observation increases. For the distribution type of L, Figure 5-7 uses the normal distribution in order for the distribution of the maximum stress σ to become a normal and symmetric distribution. The graphs in Figure 5-7 show that the area metric and PoS converge to 0.5 and one, which means the perfect separation of the prediction and observation. The result in Figure 5-7 shows that the area metric and PoS can distinguish the mean discrepancy of symmetric distributions. The parameter L in Figure 5-8 uses the lognormal distribution to produce the asymmetric response. The graphs in Figure 5-8 also present that the area metric and PoS converge to 0.5 and one. The overall result in Figure 5-7 and Figure 5-8 indicates that the area metric and PoS can represent the perfect separation between two distributions in case of the mean discrepancy.

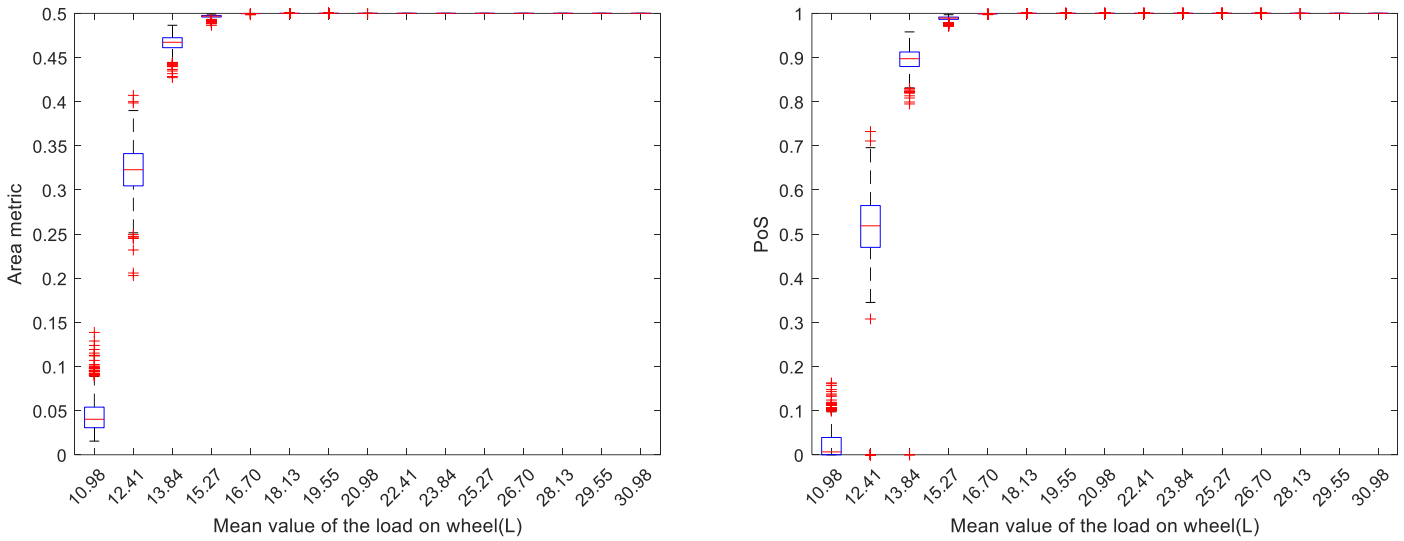
In contrast, Figure 5-9 and Figure 5-10 give different results. Figure 5-9 and Figure 5-10 consider the values of area metric and PoS when the variance of the L for the observation increases. The parameter L in Figure 5-9 uses the normal distribution, and Figure 5-10 uses the lognormal distribution. Even though the S.t.d. of the L increased to 52.10kN, more than 50 times of initial S.t.d. (e.g., 1.10kN), the values of the area metric and PoS in Figure 5-9 and Figure 5-10 do not increase to

their maximum, 0.5 and 1. In particular, the values of PoS when L follows a normal distribution, as shown in Figure 5-9 (b), rarely change. This graph supports the result in (4.15) that the limit of PoS, when the variance discrepancy approaches infinity, is zero. By the graphs in Figure 5-9 and Figure 5-10, the area metric and PoS cannot distinguish the difference of S.t.d., no matter how much the discrepancy of S.t.d. between observation and prediction increases.



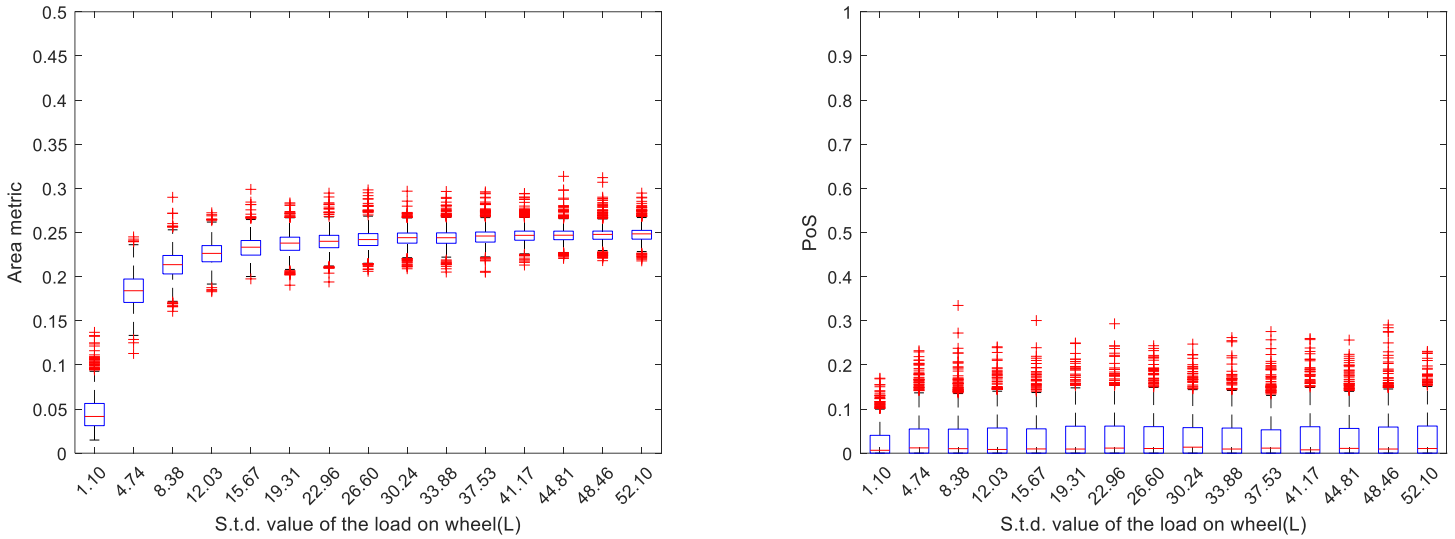
(a) (b)

Figure 5-7 The validation metrics change when the mean value of the load on wheel increases (normal distribution); (a) Area metric, (b) PoS



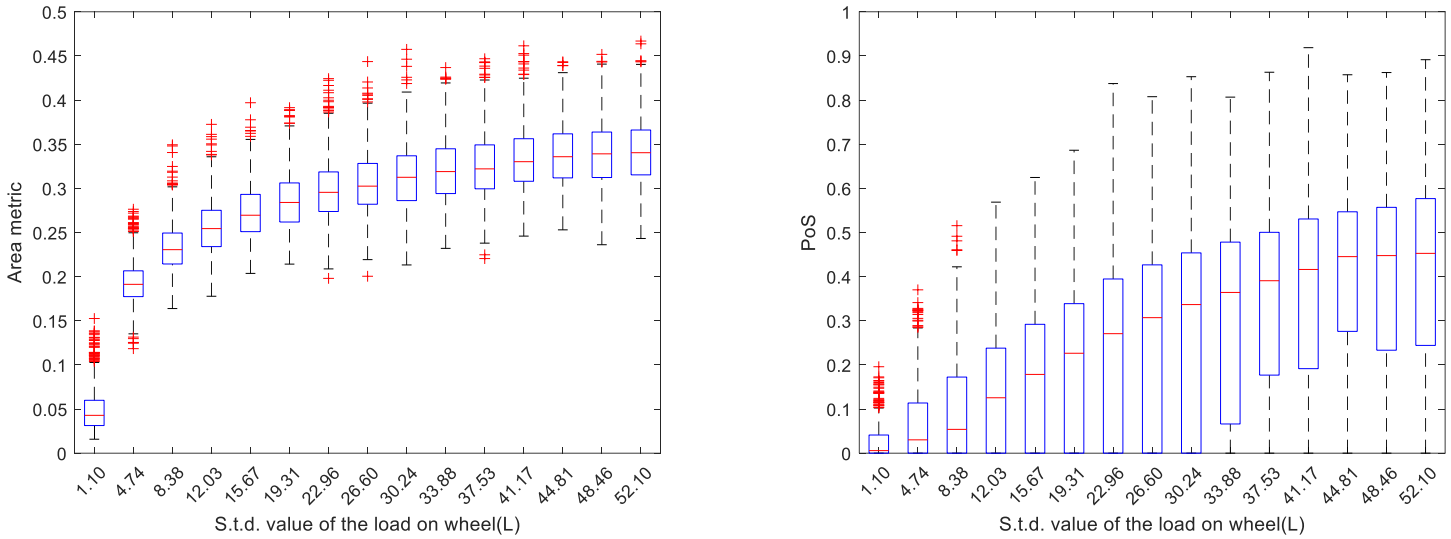
(a) (b)

Figure 5-8 The validation metrics change when the mean value of the load on wheel increases (lognormal distribution); (a) Area metric, (b) PoS



(a) (b)

Figure 5-9 The validation metrics change when the standard deviation (S.t.d.) value of the load on wheel increases (normal distribution); (a) Area metric, (b) PoS



(a) (b)

Figure 5-10 The validation metrics change when the standard deviation (S.t.d.) value of the load on wheel increases (lognormal distribution); (a) Area metric, (b) PoS

5.5 Summary and Discussion

Chapter 5 conducted a comparison study of six statistical validation metrics for their potential to offer an accurate SMV with hypothesis testing. In particular, the study focuses on the capability of these approaches to distinguish the discrepancy of variance between observed and predicted results. When the variance discrepancy is a major source of invalidity, the KLD and Bayes factor approaches give the best accuracy, among the six validation metrics studied. One noticeable thing is that the accuracy of SMV using the Bayes factor approach strongly depends on a reliable prior information. The area metric and PoS can provide a favorable accuracy when minimal observations are available, regardless of the discrepancy in the variance. This chapter provides an enhanced demonstration of why the area metric with U-pooling and PoS provides low accuracy when there is a significant discrepancy in the variances of observation and prediction. The area metric and PoS vary in a smaller range when the discrepancy in the variance increases, as compared with the discrepancy of the mean. Using automotive wheel rim model, this study emphasized that the discussion in chapters 5.3.1 and 5.3.2 is reasonable.

This is a pioneering work to evaluate statistical validation metrics under variance discrepancy in observed and predicted results. This research examines a nonlinear model that includes statistical input parameters that induce variance change, along with unknown parameters. It is also worth examining suitable statistical validation metrics for accurate SMV, since variance discrepancies between observation and prediction results commonly occur in engineering fields. I believe that the results of this study can provide a general guideline for field engineers who are not familiar with SMV, and help them to select a statistical validation metric for

hypothesis testing that leads to accurate SMV. The result is applicable for multivariate problems. However, a limitation of this study is that this study has no discussion on the problem where the difference of the mean and the variance occurs similarly. The research has not provided a quantitative criterion for selecting a statistical validation metric between the mean-favored metric (e.g., area metric and PoS) and the variance-favored metric (e.g., KLD). The future work will explore a systematic framework for SMV that can guarantee its accuracy in a variety of engineering case studies through a hybrid validation metric approach using the characteristics of validation metrics discussed in this study.

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

1) **Hyejeong Son**, Hyunhee Choi, Wongon Kim, Byeng D. Youn, and Guesuk Lee, "A Comparative Study of Statistical Validation Metrics with Consideration of Variance to Adress Type II Errors in Statistical Model Validation," *Structural and Multidisciplinary Optimization*, Submitted in April 2021.

Chapter 6

Conclusion

6.1 Contributions and Significance

The proposed research in this doctoral dissertation aims at enhancement of OBMI process with consideration of a variety of error sources in a computational model. This doctoral dissertation is composed of three research thrusts: (1) experimental design to identify error sources in optimization-based model improvement; (2) proportionate bias calibration with bound information to consider unrecognized model form errors; and (3) comparison of statistical validation metrics to reduce type II Errors in model validation. It is expected that the proposed research offers the following potential contributions and broader impacts in statistical model updating fields.

Contribution 1: New Experimental Design which Reduces Parameter Estimation Errors in Model Calibration for Error Sources Identification

Research Thrust 1 in this doctoral dissertation proposes a new experimental design

method that uses the analytical equation of parameter estimation errors. The research brought up a problem in OBMI that the process cannot deal with parameter estimation errors, coupled with model form errors. The errors in parameter estimation and model form are questionable in quantifying the error sources since the actual prediction is unknown. According to the analytical derivation of parameter estimation errors in model calibration, the new experimental design for OBMI minimizes the pseudo inverse of the gradient of model prediction, partial terms in the equation of parameter estimation errors. The method has the advantage that the experimental design can be performed only with the initial prediction model without any priors and observation data. With two case studies, the model calibration with the observation selected from the proposed methods estimates the best values for unknown parameters. This estimates can upgrade the predictability of the computational model.

Contribution 2: Novel Framework of Model Calibration to Consider the Discrepancy due to the Error Sources Using a Bound Information for Unknown Model Parameters

This doctoral dissertation devises the proportionate bias calibration method in OBMC to examine the amount of discrepancy between observation and prediction, due to the error sources such as model form and measurement errors. To the best of the authors' knowledge, most existing works in model calibration to consider the error sources are focused on how to quantify the error sources with numerous sorts of observation. By applying the proposed methods to case studies, this research

confirmed that the proposed method can contribute to: 1) adopting experience-based bound information to support the insufficient observation data for quantifying the effect of error sources, 2) preventing OBMC from optimizing the unknown parameters to the biased locations from the actual location to supplement the errors in predicted responses, and 3) being favorable for the model calibration with multiple responses whose magnitude of each response value is severely different. According to these contributions, enhanced model calibration is available.

Contribution 3: Guidance of Statistical Validation Metrics Selection for Model Validation From the Perspective of Mean and Variance Difference Between Observation and Prediction

Research Thrust 3 aims to provide guidelines and rationales to select statistical validation metrics to escape the Type II errors in model validation. To the best of the author's knowledge, the previous work only focuses on applying the validation metrics in a specific condition, such as multiple responses or lack of observation data. This research is a pioneering work to evaluate statistical validation metrics under mean and variance discrepancy in observed and predicted results. The research performed a comparison study of statistical validation metrics under the discrepancy of mean or variance between observation and prediction to confirm which metrics show the highest validation accuracy. The result reveals that area metric and PoS are sensitive to the discrepancy of the mean (Mean-supportive metrics). KLD and Bayes factor metrics are favorable to quantify the discrepancy of variance (Variance-supportive metrics). The mean-supportive metrics are advantageous in only a small

observation dataset since the quantification of variance discrepancy requires enough observation. This research has scientific merits in demonstrating why mean-supportive metrics show poor accuracy in quantifying the variance discrepancy.

6.2 Suggestions for Future Research

This doctoral dissertation performed extensive works to develop an enhanced OBMI process to deal with the error sources in computational models. Although the technical advances proposed in this doctoral dissertation successfully address some issues in OBMI, there are still several research topics that further investigations and developments are required to deliver the robust OBMI process. Specific suggestions for future research are listed as follows.

Suggestion 1: Probabilistic Experimental Design Approach for a Severely Nonlinear Response

The proposed experimental design adopts an assumption that the interested responses can be linearly simplified. Since the linear model has a constant gradient, the experimental design method gives the same result regardless of the initial value of unknown model parameters. Thus, future work should focus on a systematic method to consider a nonlinear response, which can deal with the variability of the experimental design result about the initial value of unknown model parameters. As a suggestion, the statistical information of the unknown model parameters might be

promising to assess the reliability of the experimental design result (Huan and Marzouk 2013) (Fedorov and Leonov 2013).

Suggestion 2: Consideration of Multivariate & Multiple Responses in Model Calibration

For now, the proportionate bias calibration employs a single term for all of the responses used for model calibration. The case studies considered the model calibration of two unknown model parameters so that only two sorts of observations are required. However, model calibration requires more observations as the number of unknown model parameters increases. Future research requires a discussion of whether a single-term bias can consider the effects of discrepancies due to error sources. For multiple-responses model calibration, the proportionate calibration can increase the number of bias terms. Since the dimension of the optimization space expands when the number of bias terms increases, it degrades the convergence of the optimization algorithm (Bessa et al. 2017) (Wang et al. 2020). Thus, the optimal number of bias terms should be determined.

Suggestion 3: Development of a New Validation Metric by the Integration of Mean-supportive and Variance-supportive Characteristics in Statistical Validation Metrics

The development of a validation metric that shows feasibility in the mean and

variance discrepancy simultaneously, is a significant research need. While this doctoral dissertation confirmed validation metrics that have mean-supportive and variance-supportive characteristics for a reasonable selection of metrics, most of the practical cases involve mean and variance discrepancy together. Thus, future works should focus on the integration of mean-supportive and variance-supportive characteristics in a newly formulated validation metric. In addition, the newly formulated validation metric should consider the effect of the number of observations since the mean-supportive and variance-supportive characteristics can differ according to the number of observations.

Appendix A

Analytical Derivation of Probability of Separation (PoS) with Normal and Lognormal Distribution

To find out the range of PoS when the difference of variance between observation and prediction increases, these appendices provide the derivation of the PoS using an analytical probability distribution function. Appendix A.1 assumes that the observation and prediction follow a normal distribution; Appendix A.2 assumes a lognormal distribution. The choice of a normal distribution or lognormal distribution from among the numerous distribution types is made to consider PoS when the observation and prediction follow the symmetric and asymmetric distribution.

A.1 Analytical Derivation of PoS Metric with a Normal Distribution

The general PDF of the normal distribution is as follows.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-0.5\left(\frac{x-\mu}{\sigma}\right)^2} = \varphi\left(\frac{x-\mu}{\sigma}\right) \quad \text{(A.1)}$$

where μ, and σ denote the mean and variance, which are the distribution parameters of a normal distribution. φ denotes the standard normal distribution function, which gives zero and one as a mean and standard deviation. The CDF of a normal distribution is shown in (A.2).

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-0.5\left(\frac{t-\mu}{\sigma}\right)^2} dt = \Phi\left(\frac{x-\mu}{\sigma}\right) \quad \text{(A.2)}$$

The CDF $F(x)$ is an integral of the PDF in a range of [-∞, x]. Φ stands for the function of a standard normal CDF. Using (A.1), and (A.2), the PNS in (4.7) is rewritten as follows.

$$P_{NS} = \int_{-\infty}^{\infty} F_{c_2} f_{c_1} dx = \int_{-\infty}^{\infty} \Phi\left(\frac{x-\mu_2}{\sigma_2}\right) \frac{1}{\sigma_1\sqrt{2\pi}} e^{-0.5\left(\frac{x-\mu_1}{\sigma_1}\right)^2} dx \quad \text{(A.3)}$$
$$\leq \mu_1 \quad \mu_2$$

With a normality assumption, the median and the mean of a distribution are the same. Thus, P_{NS} satisfies $\mu_1 \leq \mu_2$. For the integration of (A.3), the following transformation is adopted.

$$\begin{gathered} \frac{x-\mu_1}{\sigma_1} = t \\ x = \sigma_1 t + \mu_1 \\ dx = \sigma_1 dt \end{gathered} \tag{A.4}$$

Substituting (A.4) into (A.3), the P_{NS} is derived as follows.

$$\begin{aligned} P_{NS} &= \int_{-\infty}^{\infty} \Phi\left(\frac{x-\mu_2}{\sigma_2}\right) \frac{1}{\sigma_1\sqrt{2\pi}} e^{-0.5\left(\frac{x-\mu_1}{\sigma_1}\right)^2} dx \\ &= \int_{-\infty}^{\infty} \Phi\left(\frac{x-\mu_2}{\sigma_2}\right) \frac{1}{\sigma_1\sqrt{2\pi}} e^{-0.5(t)^2} \sigma_1 dt \\ &= \int_{-\infty}^{\infty} \Phi\left(\frac{\sigma_1 t + \mu_1 - \mu_2}{\sigma_2}\right) \varphi(t) dt \end{aligned} \tag{A.5}$$

The CDF term Φ in (A.5) is transformed to (A.6). Y is an artificial variable which follows standard normal distribution.

$$\begin{aligned} \Phi\left(\frac{\sigma_1 t + \mu_1 - \mu_2}{\sigma_2}\right) &= P\left(Y \leq \frac{\sigma_1 t + \mu_1 - \mu_2}{\sigma_2}\right) \\ &= P(\sigma_2 Y - \sigma_1 t \leq \mu_1 - \mu_2) = \Phi\left(\frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right) \end{aligned} \tag{A.6}$$

Since Y and t follow standard normal distribution, the summation of two variables ($\sigma_2 Y - \sigma_1 t$) follows N~(0, $\sigma_1^2 + \sigma_2^2$). Using the definition of the expectation, (A.7) provides the overall equation of P_{NS}.

$$P_{NS} = \int_{-\infty}^{\infty} \Phi\left(\frac{\sigma_1 t + \mu_1 - \mu_2}{\sigma_2}\right) \varphi(t) dt$$
$$= \int_{-\infty}^{\infty} \Phi\left(\frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right) \varphi(t) dt \quad \text{(A.7)}$$
$$= \Phi\left(\frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right) \int_{-\infty}^{\infty} \varphi(t) dt = \Phi\left(\frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right)$$

A.2 Analytical Derivation of PoS Metric with a Lognormal Distribution

The general PDF and CDF of a lognormal distribution is as follows.

$$f(x) = \frac{1}{s\sqrt{2\pi}} e^{-0.5\left(\frac{\ln(x)-m}{s}\right)^2} = \varphi(\frac{\ln(x) - m}{s}) \quad \text{(A.8)}$$

$$F(x) = \int_{-\infty}^{x} \frac{1}{s\sqrt{2\pi}} e^{-0.5\left(\frac{\ln(t)-m}{s}\right)^2} dt \quad \text{(A.9)}$$
$$= \Phi(\frac{\ln(x) - m}{s})$$

The distribution parameters m and s become the mean and standard deviation of ln(x). Similar to the derivation process in Appendix A.1, P_{NS} is derived as follows.

$$P_{NS} = \int_{-\infty}^{\infty} F_{c_2} f_{c_1} dx$$
$$= \int_{-\infty}^{\infty} \Phi\left(\frac{\ln(x) - m_2}{s_2}\right) \frac{1}{s_1\sqrt{2\pi}} e^{-0.5\left(\frac{\ln(x)-m_1}{s_1}\right)^2} dx \quad \text{(A.10)}$$

$$\frac{\ln(x) - m_1}{s_1} = t$$
$$x = e^{s_1 t + m_1} \quad \text{(A.11)}$$
$$\frac{1}{x} dx = s_1 dt$$

(A.12) is the result of transformation, substituting (A.11) into (A.10).

$$P_{NS} = \int_{-\infty}^{\infty} \Phi\left(\frac{\ln(x) - m_2}{s_2}\right) \frac{1}{s_1\sqrt{2\pi}} e^{-0.5\left(\frac{\ln(x)-m_1}{s_1}\right)^2} dx$$
$$= \int_{-\infty}^{\infty} \Phi\left(\frac{\ln(x) - m_2}{s_2}\right) \frac{1}{s_1\sqrt{2\pi}} e^{-0.5(t)^2} x s_1 dt \quad \text{(A.12)}$$
$$= \int_{-\infty}^{\infty} x\Phi\left(\frac{s_1 t + m_1 - m_2}{s_2}\right) \varphi(t) dt$$

The term x and φ(t) in (A.12) is reorganized as follows.

$$x\varphi(t) = e^{s_1 t + m_1} \times \frac{1}{\sqrt{2\pi}} e^{-0.5t^2} = \frac{1}{\sqrt{2\pi}} e^{-0.5t^2 + s_1 t + m_1}$$
$$= \frac{1}{\sqrt{2\pi}} e^{-0.5(t-s_1)^2 + \frac{s_1^2}{2} + m_1} = \frac{1}{\sqrt{2\pi}} e^{-0.5(t-s_1)^2} e^{\frac{s_1^2}{2} + m_1} \quad \text{(A.13)}$$
$$= e^{\frac{s_1^2}{2} + m_1} \varphi(t - s_1)$$

When substituting (t-s_1) into L, the overall process of P_{NS} derivation is described in (A.14) using (A.13).

$$
\begin{aligned}
P_{NS} &= \int_{-\infty}^{\infty} x\Phi\left(\frac{s_1 t + m_1 - m_2}{s_2}\right)\varphi(t)dt \\
&= \int_{-\infty}^{\infty} \Phi\left(\frac{s_1 t + m_1 - m_2}{s_2}\right) e^{\frac{s_1^2}{2}+m_1}\varphi(t - s_1)dt \\
&= e^{\frac{s_1^2}{2}+m_1}\int_{-\infty}^{\infty} \Phi\left(\frac{s_1 L + s_1^2 + m_1 - m_2}{s_2}\right)\varphi(L)dL \\
&= e^{\frac{s_1^2}{2}+m_1}\mathrm{E}\left(\Phi\left(Y \le \frac{s_1 L + s_1^2 + m_1 - m_2}{s_2}\right)\right) \\
&= e^{\frac{s_1^2}{2}+m_1}\Phi\left(Y \le \frac{s_1 L + s_1^2 + m_1 - m_2}{s_2}\right) \\
&= e^{\frac{s_1^2}{2}+m_1}\Phi\left(\frac{m_1 - m_2 - s_1^2}{\sqrt{s_1^2 + s_2^2}}\right)
\end{aligned}
\tag{A.14}
$$

To consider the P_{NS} concerning the mean and variance of distribution c_2, instead of the distribution parameters, the following (A.15) and (A.16) present the relation between the distribution parameters (e.g., m, and s) and the mean or standard deviation.

$$
m_2 = \ln\left(\frac{\mu_2^2}{\sqrt{\sigma_2^2 + \mu_2^2}}\right) \tag{A.15}
$$

$$
s_2 = \sqrt{\ln\left(\frac{\sigma_2^2}{\mu_2^2} + 1\right)} \tag{A.16}
$$

Using the (A.15) and (A.16), P_{NS} in (A.14) is rewritten about the mean and standard deviation.

$$P_{NS} = e^{\frac{s_1^2}{2}+m_1}\Phi\left(\frac{m_1 - \ln\left(\frac{\mu_2^2}{\sqrt{\sigma_2^2+\mu_2^2}}\right) - s_1^2}{\sqrt{s_1^2 + \ln\left(\frac{\sigma_2^2}{\mu_2^2}+1\right)}}\right) \quad \text{(A.17)}$$

References

Alcácer V, Cruz-Machado V (2019) Scanning the industry 4.0: A literature review on technologies for manufacturing systems. Int J Eng Sci Technol 22:899-919

Anderson D, Burnham KJSNS-V (2004) Model selection and multi-model inference. 63:10

Arendt PD, Apley DW, Chen W (2012) Quantification of model uncertainty: Calibration, model discrepancy, and identifiability. J MECH DESIGN 134

Arendt PD, Chen W, Apley DW Updating predictive models: Calibration, bias correction and identifiability. In: International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, 2010. pp 1089-1098

Atkinson A, Donev A, Tobias R (2007) Optimum experimental designs, with SAS vol 34. Oxford University Press, United Kingdom

Babuska I, Oden JT (2004) Verification and validation in computational engineering and science: basic concepts. COMPUT METHOD APPL M 193:4057-4066

Baig SA (2020) Bayesian Inference: An Introduction to Hypothesis Testing Using Bayes Factors. Nicotine Tob Res 22:1244-1246 doi:10.1093/ntr/ntz207

Balci O, Sargent RG (1982) Some examples of simulation model validation using hypothesis testing. Institute of Electrical and Electronics Engineers (IEEE),

Bandara S, Schlöder JP, Eils R, Bock HG, Meyer T (2009) Optimal experimental design for parameter estimation of a cell signaling model. PLoS Comput Biol 5:e1000558

Beck JL, Au S-K (2002) Bayesian updating of structural models and reliability using Markov chain Monte Carlo simulation. Journal of engineering mechanics 128:380-391

Behmanesh I, Moaveni B, Lombaert G, Papadimitriou C (2015) Hierarchical Bayesian model updating for structural identification. Mechanical Systems and Signal Processing 64:360-376

Berger JO, Mortera J (1999) Default Bayes factors for nonnested hypothesis testing. J AM STAT ASSOC 94:542-554 doi:Doi 10.2307/2670175

Bessa MA et al. (2017) A framework for data-driven analysis of materials under uncertainty: Countering the curse of dimensionality. 320:633-667

Bi S, Prabhu S, Cogan S, Atamturktur S (2017) Uncertainty Quantification Metrics with Varying Statistical Information in Model Calibration and Validation. AIAA Journal 55:3570-3583 doi:10.2514/1.J055733

Bishop CM (2006) Pattern recognition and machine learning. springer,

Bock HG, Körkel S, Schlöder JP (2013) Parameter estimation and optimum experimental design for differential equation models. In: Model based parameter estimation. Springer, United States, pp 1-30

Campolongo F, Cariboni J, Saltelli A (2007) An effective screening design for sensitivity analysis of large models. Environmental modelling & software 22:1509-1518

Chisari C, Macorini L, Amadio C, Izzuddin BA (2017) Optimal sensor placement for structural parameter identification. Structural and Multidisciplinary Optimization 55:647-662

de Aguiar PF, Bourguignon B, Khots M, Massart D, Phan-Than-Luu RJC, systems il (1995) D-optimal designs. 30:199-210

Dieter GE (1991) Engineering design: a materials and processing approach vol 2. McGraw-Hill New York,

Edwards AWF (1984) Likelihood. CUP Archive, United Kingdom

Fedorov VV, Leonov SL (2013) Optimal Design for Nonlinear Response Models. doi:10.1201/b15054

Ferson S, Oberkampf WL, Ginzburg L (2008) Model validation and predictive capability for the thermal challenge problem. COMPUT METHOD APPL M 197:2408-2430 doi:10.1016/j.cma.2007.07.030

Ferson S, Oberkampf WLJIJoR, Safety (2009) Validation of imprecise probability models. 3:3-22

Forman EH, Gass SI (2001) The analytic hierarchy process—an exposition. Operations Research 49:469-486

Frank PD, Shubin GRJJoCP (1992) A comparison of optimization-based approaches for a model computational aerodynamics design problem. 98:74-89

Frey HC, Patil SR (2002) Identification and review of sensitivity analysis methods. Risk Analysis 22:553-578

Gholizadeh S (2013) Structural optimization for frequency constraints. Metaheuristic Applications in Structures and Infrastructures:389

Gholizadeh SJMais, infrastructures (2013) Structural optimization for frequency constraints. 389

Hamby D (1994) A review of techniques for parameter sensitivity analysis of environmental models. Environmental Monitoring and Assessment 32:135-154

Hess PE, Bruchman D, Assakkaf IA, Ayyub BMJNej (2002) Uncertainties in material and geometric strength and load variables. 114:139-166

Higdon D, Nakhleh C, Gattiker J, Williams B (2008) A Bayesian calibration approach to the thermal problem. COMPUT METHOD APPL M 197:2431-2441

Hills RG, Pilch M, Dowding KJ, Red-Horse J, Paez TL, Babuška I, Tempone R (2008) Validation challenge workshop. COMPUT METHOD APPL M 197:2375-2380

Hills RG, Trucano TG (2002) Statistical Validation of Engineering and Scientific Models: A Maximum Likelihood Based Metric; TOPICAL. Sandia National Labs.,

Hills RG, Trucano TGJSNL, Albuquerque, NM, Report No. SAND99- (1999) Statistical validation of engineering and scientific models: Background.

Hu J, Zhou Q, McKeand A, Xie T, Choi S-K (2020) A model validation framework based on parameter calibration under aleatory and epistemic uncertainty. STRUCT MULTIDISCIP O 63:645-660 doi:10.1007/s00158-020-02715-z

Huan X, Marzouk YM (2013) Simulation-based optimal Bayesian experimental design for nonlinear systems. J COMPUT PHYS 232:288-317 doi:10.1016/j.jcp.2012.08.013

Iooss B, Lemaître P (2015) A review on global sensitivity analysis methods. In: Uncertainty management in simulation-optimization of complex systems. Springer, pp 101-122

Jeon BC, Jung JH, Youn BD, Kim YW, Bae YC (2015) Datum unit optimization for robustness of a journal bearing diagnosis system. Int J Precis Eng Man 16:2411-2425 doi:10.1007/s12541-015-0311-y

Jiang C, Hu Z, Liu Y, Mourelatos ZP, Gorsich D, Jayakumar PJCMiAM, Engineering (2020) A sequential calibration and validation framework for model

uncertainty quantification and reduction. 368:113172

John RS, Draper NR (1975) D-optimality for regression designs: a review. TECHNOMETRICS 17:15-23

Johnson RA, Miller I, Freund JE (2000) Probability and statistics for engineers vol 2000. Pearson Education London,

Jung BC, Park J, Oh H, Kim J, Youn BDJS, optimization m (2015) A framework of model validation and virtual product qualification with limited experimental data based on statistical inference. 51:573-583

Kass RE, Raftery AEJJotasa (1995) Bayes factors. 90:773-795

Kat CJ, Els PS (2012) Validation metric based on relative error. Math Comp Model Dyn 18:487-520 doi:10.1080/13873954.2012.663392

Kennedy MC, O'Hagan A (2001) Bayesian calibration of computer models. J R STAT SOC B 63:425-464

Keysers C, Gazzola V, Wagenmakers E-JJNn (2020) Using Bayes factor hypothesis testing in neuroscience to establish evidence of absence. Nature neuroscience 23:788-799

Kim T, Lee G, Youn BDJS, Optimization M (2019) Uncertainty characterization under measurement errors using maximum likelihood estimation: cantilever beam end-to-end UQ test problem. 59:323-333

Kim T, Youn BD (2019) Identifiability-based model decomposition for hierarchical calibration. STRUCT MULTIDISCIP O 60:1801-1811 doi:10.1007/s00158-019-02405-5

Koch K-R (2013) Parameter estimation and hypothesis testing in linear models. Springer Science & Business Media,

Kullback S (1997) Information theory and statistics. Courier Corporation,

Kullback S, Leibler RAJTaoms (1951) On information and sufficiency. 22:79-86

Lee G (2019) Sequential Optimization and Uncertainty Propagation for Optimization-Based Model Calibration. Seoul National University

Lee G, Kim W, Oh H, Youn BD, Kim NH (2019) Review of statistical model calibration and validation—from the perspective of uncertainty structures. STRUCT MULTIDISCIP O 60:1619-1644 doi:10.1007/s00158-019-02270-2

Lee G, Yi G, Youn BD (2018) Special issue: a comprehensive study on enhanced optimization-based model calibration using gradient information. STRUCT MULTIDISCIP O 57:2005-2025 doi:10.1007/s00158-018-1920-8

Li W, Chen S, Jiang Z, Apley DW, Lu Z, Chen W (2016) Integrating Bayesian calibration, bias correction, and machine learning for the 2014 Sandia Verification and Validation Challenge Problem. J. Verif. Validation Uncertainty Quantif. 1

Ling Y, Mahadevan S (2013) Quantitative model validation techniques: New insights. RELIAB ENG SYST SAFE 111:217-231 doi:10.1016/j.ress.2012.11.011

Ling Y, Mahadevan SJRE, Safety S (2013) Quantitative model validation techniques: New insights. 111:217-231

Liu Y, Chen W, Arendt P, Huang H-Z (2011) Toward a Better Understanding of Model Validation Metrics. J MECH DESIGN 133 doi:10.1115/1.4004223

Maes K, Lourens E, Van Nimmen K, Reynders E, De Roeck G, Lombaert G (2015) Design of sensor networks for instantaneous inversion of modally reduced order models in structural dynamics. Mechanical Systems and Signal Processing 52:628-644

Mahadevan S, Rebba R (2005) Validation of reliability computational models using

Bayes networks. RELIAB ENG SYST SAFE 87:223-232 doi:10.1016/j.ress.2004.05.001

Maupin KA, Swiler LP, Porter NWJJoV, Validation, Quantification U (2018) Validation metrics for deterministic and probabilistic data. 3

McFarland J, Mahadevan S, Swiler L, Giunta A Bayesian calibration of the QASPR simulation. In: Forty-eighth AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics and materials conference, Honolulu, HI, 2007.

Moon M-Y, Choi K, Cho H, Gaul N, Lamb D, Gorsich D Development of a Conservative Model Validation Approach for Reliable Analysis. In: International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, 2015. American Society of Mechanical Engineers, p V02BT03A057

Morey RD, Rouder JN (2011) Bayes factor approaches for testing interval null hypotheses. Psychol Methods 16:406-419 doi:10.1037/a0024377

Mosterman PJ, Zander J (2016) Industry 4.0 as a cyber-physical system study. SOFTW SYST MODEL 15:17-29

Multiphysics C (1998) Introduction to COMSOL multiphysics®. COMSOL Multiphysics, Burlington, MA, accessed Feb 9:2018

Myung IJJJomP (2003) Tutorial on maximum likelihood estimation. 47:90-100

Naylor TH, Finger JMJMs (1967) Verification of computer simulation models. 14:B-92-B-101

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

Oberkampf WL, Roy CJ (2010) Verification and validation in scientific computing.

Cambridge University Press, United Kingdom

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. NUCL ENG DES 238:716-743

Oberkampf WL, Trucano TG, Hirsch C (2004) Verification, validation, and predictive capability in computational engineering and physics. APPL MECH REV 57:345-384

Oh H, Choi H, Jung JH, Youn BDJS, Optimization M (2019) A robust and convex metric for unconstrained optimization in statistical model calibration—probability residual (PR). 60:1171-1187

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. STRUCT MULTIDISCIP O 54:1527-1541 doi:10.1007/s00158-016-1493-3

Oliver TA, Terejanu G, Simmons CS, Moser RDJCMiAM, Engineering (2015) Validating predictions of unobserved quantities. 283:1310-1335

Papadimitriou DI, Papadimitriou C (2015) Optimal sensor placement for the estimation of turbulence model parameters in CFD. INT J UNCERTAIN QUAN 5

Park S, Himmelblau DJTCEJ (1982) Parameter estimation and unique identifiability. 25:163-174

Peraković D, Periša M, Sente RE Information and communication technologies within industry 4.0 concept. In: Design, Simulation, Manufacturing: The Innovation Exchange, 2018. Springer, pp 127-134

Peraković D, Periša M, Zorić P Challenges and issues of ICT in Industry 4.0. In: Design, Simulation, Manufacturing: The Innovation Exchange, 2019. Springer, pp 259-269

Pérez-Cruz F Kullback-Leibler divergence estimation of continuous distributions. In: 2008 IEEE international symposium on information theory, 2008. IEEE, pp 1666-1670

Plumlee M (2017) Bayesian calibration of inexact computer models. J AM STAT ASSOC 112:1274-1285

Pukelsheim F (2006) Optimal design of experiments. Society for Industrial and Applied Mathematics, United States

Qi Q, Tao FJIA (2018) Digital twin and big data towards smart manufacturing and industry 4.0: 360 degree comparison. 6:3585-3593

Qiu N, Park C, Gao Y, Fang J, Sun G, Kim NHJJoMD (2018) Sensitivity-based parameter calibration and model validation under model error. 140

Rebba R, Mahadevan S, Huang SJRE, Safety S (2006) Validation and error estimation of computational models. 91:1390-1397

Reichert P, Schuwirth N (2012) Linking statistical bias description to multiobjective model calibration. WATER RESOUR RES 48

Ross SM (2020) Introduction to probability and statistics for engineers and scientists. Academic press,

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

Sankararaman S, Mahadevanb S (2015) Integration of model verification, validation, and calibration for uncertainty quantification in engineering systems. RELIAB ENG SYST SAFE 138:194-209 doi:10.1016/j.ress.2015.01.023

Severini TA (2000) Likelihood methods in statistics. Oxford University Press, United Kingdom

Smith A, Naik PA, Tsai C-LJJoE (2006) Markov-switching model selection using Kullback–Leibler divergence. 134:553-577

Son H, Lee G, Kang K, Kang YJ, Youn BD, Lee I, Noh Y (2020) Industrial issues and solutions to statistical model improvement: a case study of an automobile steering column. STRUCT MULTIDISCIP O 61:1739-1756 doi:10.1007/s00158-020-02526-2

Song Z, Chen Y, Sastry CR, Tas NC (2009) Optimal observation for cyber-physical systems: a fisher-information-matrix-based approach. Springer Science & Business Media,

Stanton A, Wiegand D, Stanton G (2000) Probability reliability and statistical methods in engineering design.

Sun N-Z, Sun A (2015) Model calibration and parameter estimation: for environmental and water resource systems. Springer, United States

Tabatabaian M (2015) COMSOL5 for Engineers. Stylus Publishing, LLC,

Thakur A, Banerjee AG, Gupta SKJC-AD (2009) A survey of CAD model simplification techniques for physics-based simulation applications. 41:65-80

Thonhofer E, Luchini E, Kuhn A, Jakubek S Online parameter estimation for a flexible, adaptive traffic network simulation. In: 2014 International Conference on Connected Vehicles and Expo (ICCVE), 2014. IEEE, pp 937-938

Tricaud C, Dariusz MP, Yang U, Chen Q D-optimal trajectory design of heterogeneous mobile sensors for parameter estimation of distributed

systems. In: 2008 American Control Conference, 2008. IEEE, pp 663-668

Trucano TG, Swiler LP, Igusa T, Oberkampf WL, Pilch M (2006) Calibration, validation, and sensitivity analysis: What's what. RELIAB ENG SYST SAFE 91:1331-1357

Ucinski D (2004) Optimal measurement methods for distributed parameter system identification. CRC press, United States

Wang Y, Yue X, Tuo R, Hunt JH, Shi JJAoAS (2020) Effective model calibration via sensible variable identification and adjustment with application to composite fuselage simulation. 14:1759-1776

Weathers J, Luck R, Weathers J (2009) An exercise in model validation: Comparing univariate statistics and Monte Carlo-based multivariate statistics. RELIAB ENG SYST SAFE 94:1695-1702

White L, West T (2019) Area Validation Metric for Applications with Mixed Uncertainty.

Wilcox RR (2011) Introduction to robust estimation and hypothesis testing. Academic press,

Xi Z, Fu Y, Yang R (2013) Model bias characterization in the design space under uncertainty. International Journal of Performability Engineering 9:433-444

Xiong Y, Chen W, Tsui K-L, Apley DW (2009) A better understanding of model updating strategies in validating engineering models. COMPUT METHOD APPL M 198:1327-1337

Youn BD, Jung BC, Xi Z, Kim SB, Lee W (2011) A hierarchical framework for statistical model calibration in engineering product development. COMPUT METHOD APPL M 200:1421-1431

Youn BD, Xi Z, Wang P (2008) Eigenvector dimension reduction (EDR) method for

sensitivity-free probability analysis. STRUCT MULTIDISCIP O 37:13-28 doi:10.1007/s00158-007-0210-7

Zhao LF, Lu ZZ, Yun WY, Wang WJ (2017) Validation metric based on Mahalanobis distance for models with multiple correlated responses. RELIAB ENG SYST SAFE 159:80-89 doi:10.1016/j.ress.2016.10.016

국문 초록

컴퓨터 모델 내 오류 원인 식별을 위한 최적화 기반 모델 개선 기법 연구

서울대학교 대학원
기계항공공학부
손 혜 정

컴퓨터 이용 공학 기술의 활용도가 증가함에 따라, 각 공학분야에서는 보다 정확한 예측 능력을 가진 컴퓨터 모델을 필요로 하게 되었다. 많은 연구결과를 통해, 신뢰도 높은 계산모델을 얻기 위한 공학기술들이 개발되었다. 최적화 기반 모델 향상 기술은 계산모델 예측도 향상을 위한 공학기술 중 하나로, 모델 보정, 모델 검증, 그리고 모델 개선 과정을 포함하고 있다. 모델 보정은 계산 모델 내 미지변수의 값을 역으로 추정하는 기술이다. 모델 검증은 예측 성능의 정확도를 판단한다. 계산모델 내 미지 오류 원인이 존재하면 모델 개선을 통해 미지 원인을 탐색하는 작업을 수행한다. 최적화 기반 모델향상기술 내 세가지 세부 기술들은 모델 관련 사전 정보의 양에 따라 유기적으로, 혹은 개별적으로도 수행이 가능하다.

모델 향상 기술이 계산모델 내 영향을 주는 다양한 오류원인을 고려하여 수행되고 있으나, 최적화 기반 모델 향상기술은 여전히 계산모델의 정확도를 증가시키는데 한계점을 지니고 있다. 시험 데이터 및 계산 모델 내 다양한 오류 소스들이 결합되어 있어, 최적화 기반 모델 향상 기술은 이 오류원인들을 구분하고 각 오류원인들에 대해 적합한 솔루션을 제공하기에 부적합하다. 따라서, 이러한 문제점을 해결하고자 본 박사학위논문에서는 (1) 파라미터 추정 오류 감소를 위한 시험 설계 기법, (2) 모델 보정 시 모델링 및 시험 오류의 양을 정량화 하기 위한 비율 편향도 정량화 기법 (3) 2종 오류에 강건한 통계기반 검증 척도 비교 연구를 제안하고자 한다.

첫 번째 연구에서는 파라미터 추정 오류를 최소화하기 위한 시험 설계법 개발을 목표로 한다. 여기서 결정된 시험설계안은 모델 보정 시 사용될 시험 데이터 취득을 위한 시험 설계를 뜻한다. 계산모델 내 발생하는 모델링 오류, 그리고 시험데이터 취득 시 발생하는 계측오류 등은 모델 보정에서 정확한 파라미터 값의 추정을 방해한다. 파라미터 추정 오류를 포함한 계산모델은 주어진 시험데이터를 잘 모사하는 것처럼 보이지만, 파라미터를 과도하게 편향된 값으로 추정하여 모델링 오류를 보완한 결과이다. 이 경우, 모델 검증 시 모델이 유효하다고 판단될 수 있지만 실제로는 파라미터 추정오류와 모델링 오류를 동시에 갖고 있으므로 다양한 설계조건에서 유효하지 않은 모델이다. 따라서, 본 연구에서는 파라미터 추정오류와 모델링 오류를 구분하여 정확한 모델 검증을 유도하고자 한다. 파라미터 추정오류와 모델링 오류는 그 정도를 각각 정량화 하는 것이 불가능하므로, 파라미터 추정오류를 가장 최소화 할 수 있는 시험데이터의 종류와 취득위치를 선정할 수 있는 시험설계법을 고안하였다. 이를 위해, (1) 파라미터 추정오류를

수식적으로 유도하였고, (2) 유도된 식 내에서 사용자가 제어할 수 있는 일부항을 최소화 하도록 하였다. 제안된 시험설계법은 파라미터 추정오류와 모델링 오류를 구분하고, 모델 검증 시 유효 및 불유효의 원인이 모델링오류가 될 수 있도록 한다.

두 번째 연구에서는 파라미터 추정오류를 개선하기 위해 모델 보정 시 모델링 오류에 의한 성능 저하량을 정량화 할 수 있는 비율 편향 보정 기법을 제안한다. 첫 번째 연구에서 제안한 시험설계법은 별도의 추가 시험 데이터 없이 파라미터 추정 오류와 모델링 오류를 구분해 낼 수 있는 최선의 방법론 이지만, 모델링 오류 및 시험 오류의 영향이 큰 경우 파라미터 추정오류를 획기적으로 개선하는데 한계가 있다. 오류의 영향도가 큰 모델은 추정 파라미터의 값이 엔지니어가 가진 경험, 혹은 물리 기반 정보에 위배되는 지점으로 수렴할 수 있다. 따라서, 본 연구에서는 관측데이터 외 미지 모델 변수의 물리적 정보를 활용하여 모델링 오류 및 관측오류에 의한 성능저하도의 양을 정량화 하고자 한다. 연구에서 제안된 '비율편향' 은 오류에 의한 성능저하도를 성능값의 일정한 비율로 가정하여, 모델 보정 시 최적화 알고리즘 내에서 미지모델변수와 함께 최적 값이 추정되는 항이다. 비율편향 항과 미지모델 변수가 사전의 물리적 정보에 위배되지 않는 범위 내에서 추정될 수 있도록 미지모델 변수의 범위 정보를 최적화 알고리즘의 제한조건으로 활용한다. 비율편향 보정기법은 미지모델변수의 추정값이 모델링 오류에 의한 성능저하를 보완하기 위해 과도하게 편향된 값으로 최적화 되는 현상을 바로잡을 수 있다.

세 번째 연구에서는 모델 검증 시 발생할 수 있는 결정 오류를 개선하기 위해 통계적 검증 척도의 선택 기준을 제시하고자 한다. 모델 검증은 주로 통계기반 방법인 가설검증을 활용하여 모델의 유효 및

불유효를 결정한다. 가설검증은 제 1종 오류 및 제 2종 오류의 발생 가능성을 갖고 있다. 제 2종 오류는 불유효한 모델을 유효하다고 판단하는 오류로써 실제 산업분야에 치명적인 사고를 유발할 수 있다. 본 연구에서는 제 2종 오류를 가장 적게 발생 시킬 수 있는 통계적 검증 척도를 분석하기 위해 다음과 같은 조건에서의 검증 정확도 비교 연구를 수행한다. 1) 관측 및 예측 성능의 분산이 같고 평균값의 차이로 인해 예측 성능이 불유효 한 경우, 2) 관측 및 예측성능의 평균보다 분산값의 차이로 인해 예측 성능이 불유효 한 경우. 비교연구는 모델 파라미터의 분산 정도를 4가지로 세분화 하고 관측 데이터 개수에 의한 정확도 차이를 비교하고자 관측 데이터를 3개에서 30개까지 증가시켰다. 그 결과, 성능 간 평균의 차이를 잘 정량화 하는 검증척도 및 성능 간 분산의 차이를 잘 정량화 하는 검증척도를 제안할 수 있었다. 제안된 검증척도의 평균지향 및 분산지향 특성을 증명하고자, 평균지향 척도의 극한값을 유도하여 분산값의 증가 시 척도의 값이 최대값에 도달하지 않아 검증 오류가 발생할 수 있음을 확인하였다.

주요어: 최적화 기반 모델 향상 기술
오류 원인 식별
행렬식 기반 시험 설계
비율 편향 모델 보정
통계적 검증 척도
가설 검증
디지털 트윈

학 번: 2015-20733