Final Report for Period: 09/2008 - 08/2009 **Submitted on:** 12/02/2009

Principal Investigator: Tsui, Kwok-Leung . **Award ID:** 0522366

Organization: GA Tech Res Corp - GIT

Submitted By:

Tsui, Kwok-Leung - Principal Investigator

Title:

Collaborative Research: Validating Predictive Models in Engineering Design

Project Participants

Senior Personnel

Name: Tsui, Kwok-Leung

Worked for more than 160 Hours: Yes

Contribution to Project:

Post-doc

Graduate Student

Name: Wang, Shuchun

Worked for more than 160 Hours: Yes

Contribution to Project:

Shuchun is the ph.d. student who works on the project as a research assistant. She helped develop a new Bayesian procedure for validating computer models. The developed method has been tested and compared in real examples. The research results have been written in research papers and included in her pd.d. thesis.

Name: Liu, Xuyuan

Worked for more than 160 Hours: Yes

Contribution to Project:

X. Liu has worked on the project by developing modeling and validation methods for computer experiments with functional output. He has included the results in the conference proceeding paper as well as parts of his phd thesis.

Name: Han, Sung Won

Worked for more than 160 Hours: Yes

Contribution to Project:

S.W. Han has worked on developing computation programs for implementating the proposed modeling methods. The programs have been utilized in simulation and numerical studies for comparing proposed methods with existing methods. THe results of these studies have bee reported in conference proceeding and journal papers.

Undergraduate Student

Technician, Programmer

Other Participant

Research Experience for Undergraduates

Organizational Partners

Northwestern University at Chicago

Prof. Wei Chen, Mechanical Engineering, has been working with the PI in this collaborative research. The deatiled contribution is described in the 'research activities'.

General Motors Corporation

We are working with the GM Global Performance Integration Group on a computer model validation project with profile measurement.

Other Collaborators or Contacts

Activities and Findings

Research and Education Activities: (See PDF version submitted by PI at the end of the report)

This collaborative research represents the joint efforts from two universities, Northwestern University (NU) and Georgia Tech (GT), for the development of a model validation approach that provides quantitative assessments of uncertainty in using predictive models in engineering design. During the first year of this project, the focus has been on implementing the following two research tasks: (1) Development of a Bayesian approach to assess the uncertainty in model prediction by combining data from both physical experiments and the computer model; and (2) Development of a design-oriented model validation metric to guide designers for achieving high confidence of using predictive models in making design decision. The two teams at NU and GT have worked very closely on the above two tasks while each team takes the lead on one subject, i.e., GT for task (1) and NU for task (2). The research results have been documented in two joint publications [1, 2], which are highlighted in the attached file on 'research findings'.

During the second year of this project, the two teams focused on two specific research tasks, (1) Develop a sequential design procedure to sequentially choose testing sites for running physical experiments to meet pre-specified prediction accuracy requirement; (2) Develop validation procedures for validating computer models with profile (functional) outputs. The two teams at NU and GT have worked very closely on the above two tasks while each team takes the lead on one subject, i.e., NU for task (1) and GT for task (2). The research results have been documented in two joint publications [4, 5, 6], which are highlighted in the attached file on 'research findings'.

During the third year and the extended year of the project, we have focused on developing new modeling tools for more complex validation problems with functional response data. We have developed a maximum likelihood estimation approach for modeling the combined data from physical experiments and computer model outputs. In addition, we developed a regression based approach for modeling and validating the functional output from computer models and actual physical experiments. An automotive transmission example was used to investigat and test the feasibility of the developed approach. The results have been reported and published in two joint publications, which are highlighted in the former attached file on 'research findings'.

Regarding the education activities, the research results have directly benefited the teaching of ISyE 6413 - Design and Analysis of Experiments, ISyE 6414 - Regression Analysis and Statistical Modeling, ISyE 7400 - Advanced Design of Experiments, and ISyE 7416 - Data Mining and Statistical Learning, all graduate level courses. Model validation and analysis methods are some of the new topics added to these courses.

Findings: (See PDF version submitted by PI at the end of the report)

This collaborative research represents the joint efforts from two universities, Northwestern University (NU) and Georgia Tech (GT), for the development of a model validation approach that provides quantitative assessments of uncertainty in using predictive models in engineering design. During the past two years of this project, the focus has been on implementing the following four research tasks: (1) Development of a Bayesian approach to assess the uncertainty in model prediction by combining data from both physical experiments and computer model; (2) Development of a design-oriented model validation metric to guide designers for achieving high confidence of using predictive models in making design decision; (3) Development of objective-oriented sequential experiment design strategy; and (4) Modeling and validating computer models with functional responses. The two teams at NU and GT have worked very closely on all tasks. The research results have been documented in six joint publications.

(1) A Bayesian Approach to Model Uncertainty Quantification
Most research in validating computer models had focused on estimating prediction bias and improving accuracy of a computer model. Much less work had been done on characterizing prediction uncertainty and prediction bias under general situations. In this research we have developed a Bayesian approach to assess the uncertainty in model prediction by combining data from both physical experiments and computer model outputs, which will provide more accurate prediction than the existing methods. We also develop a step-by-step procedure for validating the computer model based on the prediction from the Bayesian approach.

(2)Design Validation Metric
Different from the existing validation metrics that assess the predictive capability (accuracy) of a model, the design validation metrics MD has been developed in our work to provide a probabilistic measure of whether the real outcome of a candidate design is better than other design choices. Such metric is developed to provide a direct measure of how reliable is the decision of choosing one design candidate versus the other design alternatives, therefore to provide the confidence associated with a design decision with the consideration of model uncertainty. In addition, such metric provides useful guidance for validation activities.

(3) Sequential Experiment Design
A sequential sampling strategy is first developed for computer experiments in variable fidelity optimization [3]. We applied the Bayesian approach developed in the model validation research to model fusion for integrating high fidelity (HF) and low fidelity (LF) models into the predicative surrogate model, over which design optimizations are performed. The developed sequential sampling strategy is intended to overcome existing sequential sampling methods.

(4) Computer Validation with Functional Responses
Previous work on computer validation had been focused on problems with single output (response). In reality, performance output of a system can be represented as a profile (functional form), such as a response over time. Our research team has worked with the General Motors Global Performance Integration Group on a project on modeling and validating computer models with functional responses. In particular, we have focused on a computer model on simulating the profile of acceleration over time at different input conditions of tire coefficient, drag coefficient, road grade, and throttle position. An experiment of 64 runs (two replications at each of 32 combinations of input conditions) have been conducted to collect the acceleration profiles of both the physical experiments and computer simulation outputs. The objectives are to build reliable empirical models to (i) predict acceleration profile at untested conditions, (ii) predict ending acceleration and shift time, and (iii) validate the computer model with physical experiments.

Training and Development:

Graduate students supported under this grant had the opportunity to learn how to conduct collaborative research with researchers from a different research institution and with different background. Students were exposed to various issues related to probability and statistical analyses, engineering design, uncertainty modeling, etc. The project also provides the learning opportunity of presenting and publishing research results. Shuchun Wang, who worked on the project in the first year, had graduated and received her Ph.D. in summer of 2006. Xuyuan Liu, a Ph.D. student, was working on this project in 2007 and 2008. Sung Won Han, another Ph.D. student was working on the project in 2009.

Outreach Activities:

In the course of this project, the research team has exchanged research ideas with many other research groups that have similar interests in the topic of model validation, which helps identify the research needs of the proposed project. Examples of these research groups include the Optimization and Uncertainty Estimation group at the Sandia National Laboratory, the design methodology group at Ecole Central Paris, the Safety Engineering group in the Scientific Research Lab of Ford Motor Company, and the Global Performance Integration Group at General Motors. Research results have been presented at the Stochastic Modeling workshop at University of Notre Dame (March 24-26, 06), the Panel Session on 'Transition of Non Deterministic Approaches from Academic and National Lab Research to Industrial Design and Decision-Making', at the SAE congress (April 6th, 06), the 2006, 2007, 2008, 2009 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference., the 7th World Congress on Structural and Multidisciplinary Optimization, the 2008 49th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, the 2006, 2007, 2008 INFORMS Annual Conference, and the First, Second, and Third Pre-Conference Workshop of Data Mining. Professors Chen and Tsui have also delivered invited talks on the research subject at a number of university seminars (e.g., Purdue University, University of Florida, University of Texas-Arlington, Virginia Tech, Chinese University of Hong Kong, City University of Hong Kong, Hong Kong University of Science and Technology, Shanghai Jiao Tong University, Tong Ji University, University of Electronic Science and Technology of China) and industry visits (e.g., Ford Motor, General Motors, Boeing, and General Electric).

Journal Publications

Chen, W., Tsui, K-L., Xiong, Y., and Wang, S., "A Design Driven Validation Approach Using Bayesian Predictive Models", ASME Journal of Mechanical Design, p. , vol. 130, (2008). Published,

Wang, S., Chen, W., and Tsui, K, "Bayesian Validation of Computer Models", Technometrics, p. , vol. , (2009). Accepted,

Xiong, Y., Chen, W., Tsui, K.-L., "A New Variable Fidelity Optimization Framework Based on Model Fusion and Objective-Oriented Sequential Sampling", Journal of Mechanical Design, p. , vol. 130, (2008). Published,

Xiong, Y., Chen, W., Tsui, K-L., and Apley, D., "A Better Understanding of Model Updating Strategies in Validating Engineering Models", Journal of Computer Methods in Applied Mechanics and Engineering, p. 1327, vol. 198, (2009). Published,

Books or Other One-time Publications

Chen, W., Tsui, K.-L., Xiong, Y., Wang, S., "Metric and a Bayesian Procedure for Validating Predictive Models in Engineering Design", (2006). Book, Accepted
Bibliography: Sept. 10-13, Philadelphia, PA, conerence paper 2006 ASME Design Technical Conference. Design Automation Conference

S. Wang, "Data Mining, Forecasting, and Computer Model Validation", (2006). Thesis, Submitted
Bibliography: School of ISyE, Georgia Tech

Xiong, Y., Chen, W., Tsui, K-L, "A New Variable Fidelity Optimization Framework Based on Model Fusion and Objective-Oriented Sequential Sampling", (2007). Conference Proceeding, Accepted
Bibliography: ASME Design Engineering Technical Conference, Design Automation Conference, September 4-7, Las Vegas, Nevada, 2007

Liu, X., Chen, W., Tsui, K.-L., "Regression Modeling for Computer Model Validation with Functional Responses", (2008). Conference paper, Published
Bibliography: Paper No. DETC2008-49662, Proceedings of the ASME 2008 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference, August 3-6, 200

Web/Internet Site

Other Specific Products

Contributions

Contributions within Discipline:

Compared to the existing work, our work focuses on a Bayesian model for predicting computer model bias and true model output, that are accurate, flexible and economically sound. In engineering applications where it is too expensive to obtain experimental data, the Bayesian inference approach offers much flexibility as additional design knowledge and information can be easily incorporated through prior distributions. With the Bayesian approach, uncertainty in prediction related to the lack of experiment data can be captured by the magnitude of uncertainty of the bias function, which offers rigorous and flexible methods for quantifying the model uncertainty in an intended design domain that may interpolate as well as extrapolate from a validation domain. Since we have developed the analytical results in implementing the Bayesian approach, the Bayesian approach we proposed can be economically implemented in multidimensional problems.

Based on the Bayesian approach we proposed, our research is the first work that provides theoretical discussion on the significance of combining computer outputs and physical, which can improve the prediction of the real system output over using only computer outputs or only physical observations.
Our research is one of the pioneering works that provide quantitative means to define and to assess model validity from the perspective of design decision making with the consideration of various sources of uncertainties. It offers a new and improved way of viewing model validation by relating its definition to a specific design choice. The proposed metric for assessing design validity provide probabilistic measurements with regard to the confidence of using a model for making a specific design choice; they can be used to overcome the limitations of many existing model validation approaches while providing direct estimate of the global impact of uncertainty sources on the confidence in a design decision.
Our research also clarified the role that model validation should play in decision making under uncertainty and developed strategies for making tradeoffs based on both product design and model validation. Unlike most of the existing model validation works that focus on the assessment of model accuracy, model validation in our research is viewed as a process to improve designer's confidence in making a design choice using the improved predictive model, which is the augmented model that includes both the original computer model as well as the estimate of the bias function. The research directly addressed the needs of Engineering Design programs that seek improvement on validation of models, increased emphasis on treatment of uncertainty, and improvement on computational tools needed to implement the theory.
Our research proposed a new and effective strategy of sequential experiment design in variable fidelity optimization, which is immediately extendable to the model validation. Using decision validation metrics for assessing the confidence of the optimum design, we are able to enhance the predictive capability of a computer model for the purpose of design decision making. By treating model uncertainty separately from design variable/parameter uncertainty, we are able to effectively design the physical experiments, to sequentially eliminate the model uncertainty.
There is a growing recognition that a model needs to be updated to better reflect the physical experiment observations that are collected in model validation. Our research provides a better understanding of the various model updating strategies, which utilize mathematical means to update a computer model based on both physical and computer observations. The Maximum Likelihood Estimation (MLE) method proposed provides a better interpretation of the observed dispersion of experimental data. Uncertainty in model prediction is quantified to account for various sources of uncertainty in a validation process. Since our approach is applied to the widely used benchmark thermal challenge problem, other researchers who are interested in this topic can further compare our results with those from their studies. The research provides more insights into the benefits and limitations of using the MLE method versus the Bayesian approach. Insights into various model updating strategies are also obtained through this study and can serve as the guideline in engineering practice.
Following the classical nonparametric regression framework, our proposed method for modeling and validating functional response uses a single step procedure which is shown to be easily implemented and computationally efficient.

Contributions to Other Disciplines:

Our research has offered a generic model validation approach that can be applied to many domestic and military applications for making reliable decisions when using predictive models as a replacement of expensive physical part deployment. Our research has leveraged the results from existing model validation work in the computational modeling community and extended their use in engineering design. Results are broadly disseminated throughout mechanical engineering, industrial engineering, simulation, and applied statistics communities. The strong collaborations between the research teams, industrial partners, and government agencies has ensured that the technology is transferred and the results are successfully implemented. The research has contributed to education in the areas of model-based simulation, modeling and optimization of engineering systems under uncertainty, statistical analysis, engineering design, and information technology as well as provide training to minority and women engineering students.

Contributions to Human Resource Development:

Three graduate students have benefited from working on the project through literature search, methodology development, and computer programming. S. Wang, X. Liu, and SW Han have received support from this project. Wang's thesis was on Bayesian Validation of Computer Models and has graduated in 2007. Liu's thesis is on Bayesian Modeling of Computer Experiments with Functional Output, and is expected to finish in the summer of 2010. Han's thesis is on a different topic but has contributed to developing software and computer programs for the research methods in this project.

Contributions to Resources for Research and Education:

Graduate students supported under this grant had the opportunity to learn how to conduct collaborative research with researchers from a different research institution and with different background. Students were exposed to various issues related to probability and statistical analyses, engineering design, uncertainty modeling, etc. The project also provides the learning opportunity of presenting and publishing research results.

The research results have directly benefited the teaching of ISyE 6413 - Design and Analysis of Experiments, ISyE 6414 - Regression Analysis and Statistical Modeling, ISyE 7400 - Advanced Design of Experiments, and ISyE 7416 - Data Mining and Statistical Learning, all graduate level courses. Model validation and analysis methods are some of the new topics added to these courses.

Contributions Beyond Science and Engineering:

Conference Proceedings

Categories for which nothing is reported:

Any Web/Internet Site
Any Product
Contributions: To Any Beyond Science and Engineering
Any Conference

Collaborative Research: Validating Predictive Models in Engineering Design

Wei Chen
Northwestern University

Kwok Tsui
Georgia Institute of Technology

Abstract

This collaborative research represents the joint efforts from two universities, Northwestern University (NU) and Georgia Tech (GT), for the development of a model validation approach that provides quantitative assessments of uncertainty in using predictive models in engineering design. During the past year of this project, the focus has been on implementing the following two research tasks: (1) Achieving a better Understanding of Model Updating Strategies in Validating Engineering Models; and (2) Modeling and validating computer models with functional responses. The two teams at NU and GT have worked very closely on all tasks. The research results have been documented in two conference papers, both are under preparation for joint publication..

Research Findings

1. A better Understanding of Model Updating Strategies in Validating Engineering Models

In this task, we examine various model updating strategies as an integral part of the model validation process. The existing model updating strategies differ in their formulations, the solution method used, and the physical interpretations. The two most widely used categories of formulations include bias-correction and calibration.

Through our examination, we found there are several limitations when applying the traditional Bayesian calibration approaches to update a computer model using either bias-correction, calibration, or a combination of both. Besides the numerical difficulty in implementation, one major limitation of the Bayesian approach is that the calibration parameters are treated as uncertain due to lack of knowledge, not accounting for sources of variability in a validation process. Besides, the choice of the prior distribution in Bayesian analysis is often arbitrary.

As an alternative approach to the traditional Bayesian approach, we examine in this task a new model updating strategy, in which a computer model is updated to better interpret the observed *dispersion* of experimental data. The Maximum Likelihood Estimation (MLE) method is used to estimate the model updating parameters as shown in Figure 1. Unlike the traditional Bayesian approach which accounts for the experimental uncertainty by a single error term, the MLE based model updating approach accounts for experimental uncertainty through a subset of model updating parameters. In contrast to the Bayesian approach, the MLE based approach does not rely on the prior distributions of calibration parameters, instead, it seeks optimal distribution parameters underlying model updating parameters through maximizing the likelihood function based on the physical experiment data (Figure 2).

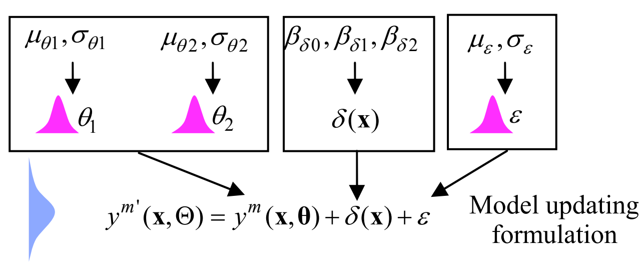

Figure 1. Model updating parameters Θ in formulation $y^{m'}(\mathbf{x},\Theta)$

Through the thermal challenge example, we demonstrate that model updating can be treated as an integral part of a model validation process which improves a model based on the physical observations gathered. We illustrate that without running into numerical complexity, the model updating method we proposed is easier to implement and interpret compared to the existing Bayesian methods.

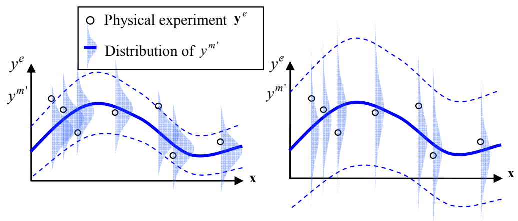

Figure 2. Likelihood value indicates the agreement between the output distribution of the updated model and the dispersion of physical experiments

Using the newly developed u-pooling method by Ferson et al, we show that the metric can be applied to both the original and the updated models to assess the accuracy and predictive capability of different model updating formulations. Through in-sample and out-sample tests (Figure 3) based on different data sets, we find that the proposed model updating approach improves the agreement between the model and the physical experiment data. However, when applying the updated model at a region that is far from the domain of data used for model updating, the extrapolation capability of the updated model is not guaranteed.

By comparing our approach to the existing works on the thermal challenge problem, we observe the differences of various methods in utilizing available data, the model updating formulations adopted, and the solution method employed. Even though our method is different, we find the conclusion we reach on device failure probability is identical to other methods in literature. As for which model updating formulation is the most appropriate, unless it can be specified based on the pre-existing knowledge, we think it is problem dependent and should be selected by exercising the model validation metrics as demonstrated.

While model updating is shown to be useful for improving the accuracy of a model, as the process is fully data-driven, we believe the method should be used with caution when used for extrapolation.

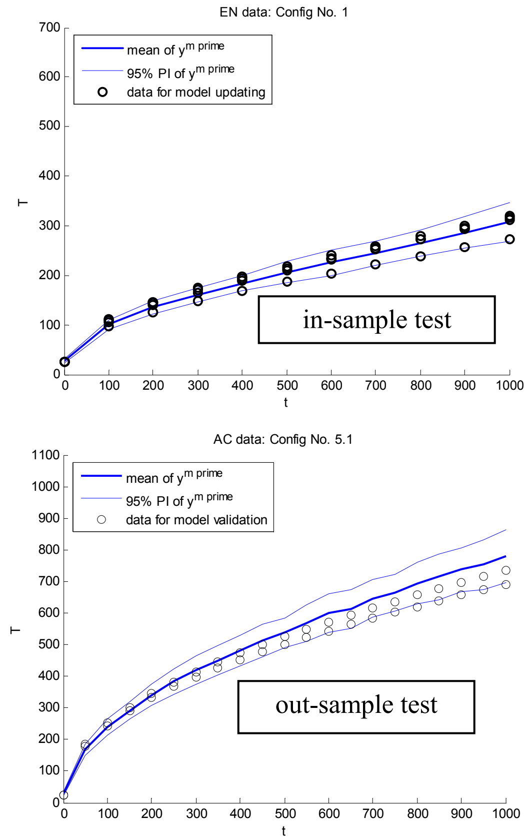

Figure 3. Example of In-Sample and Out-Sample Test

Due to the nature of the MLE method, the effectiveness and accuracy of the MLE based model updating approach could be downgraded when data amount is extremely small. In our test with the 'low level' data sufficiency for the thermal challenge problem, it is found that the bandwidth of the prediction uncertainty could be degenerated to fairly small values, unable to reflect the condition of lack of data. To mitigate this problem, prior knowledge may be used to specify more conservative bounds of model updating parameters to prevent them from running into 'absurd' values. Another potential weakness of the MLE based model updating approach might be associated with the numerical instability during the optimization of the likelihood function, especially when a complex model updating formulation that involves many parameters is considered. To mitigate this issue, sensitivity analysis could be performed prior to optimization, by leaving out parameters that are insensitive to model output and the likelihood function.

(2) Modeling and validating computer models with functional responses.

Statistical analysis of functional responses based on functional data from both computer and physical experiments has gained increasing attention due to the dynamic nature of many engineering systems. However, the complexity and huge amount of functional data bring many difficulties to apply traditional or existing methodologies. The objective of the present study is twofold: (1) prediction of functional responses based on functional data and (2) prediction of bias function for validation of a computer model that predicts functional responses.

A single step functional regression modeling approach is developed under this task to analyze functional outputs of physical and computer experiments. Traditional methods for modeling functional data generally involve two steps. Models are first fit at each individual setting of the input to reduce the dimensionality of the functional data. Then the estimated parameters of the models are treated as new responses, which are further modeled for prediction. Alternatively, pointwise models are first constructed at each time point and then functional curves are fit to the parameter estimates obtained from the fitted models. We propose a single model to relate the functional response to both the input and the time variables. To overcome the high correlation between the shift (ending) response and the shift time, a sequential procedure is proposed to model the shift time as a function of the inputs and the shift response. We find the proposed model may be easier to interpret and implement for certain applications. Through a comparison with the existing Gaussian process modeling approach using a real industrial example provided by General Motor, we demonstrate that the proposed method yields sufficient accuracy, performs efficiently and achieves satisfactory accuracy in global prediction.

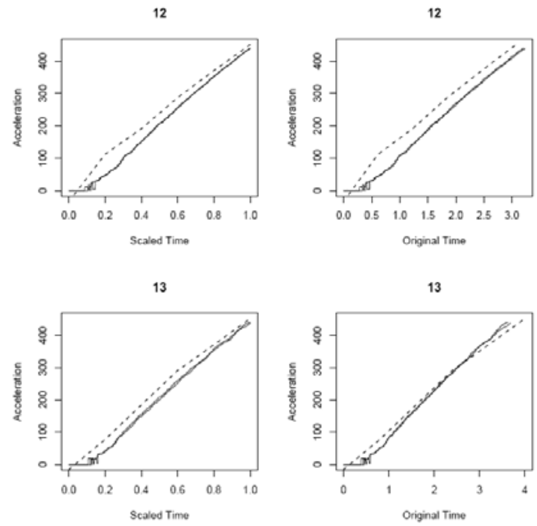

Figure 4: Predictions of two untested conditions: Observed (solid line), Predicted (dashed line)

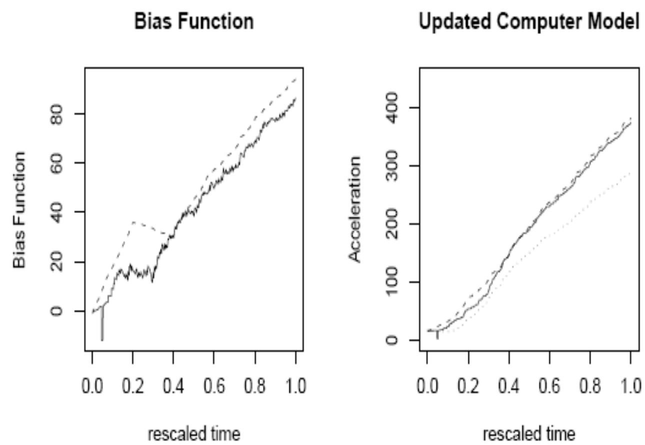

Figure 5: Bias Function and Updated Computer Outputs (Prediction: Dashed; Physical (or real bias function): Solid; Computer: Dotted)

Summary of Contributions

1. Intellectual Contributions

There is a growing recognition that a model needs to be updated to better reflect the physical experiment observations that are collected in model validation. Our research provides a better understanding of the various model updating strategies, which utilize mathematical means to update a computer model based on both physical and computer observations. The Maximum Likelihood Estimation (MLE) method proposed provides a better interpretation of the observed dispersion of experimental data. Uncertainty in model prediction is quantified to account for various sources of uncertainty in a validation process. Since our approach is applied to the widely used benchmark thermal challenge problem, other researchers who are interested in this topic can further compare our results with those from their studies. The research provides more insights into the benefits and limitations of using the MLE method versus the Bayesian approach. Insights into various model updating strategies are also obtained through this study and can serve as the guideline in engineering practice.

Following the classical nonparametric regression framework, our proposed method for modeling and validating functional response uses a single step procedure which is shown to be easily implemented and computationally efficient.

3. Significance (Impact)

Our research has offered a generic model validation approach that can be applied to many domestic and military applications for making reliable decisions when using predictive models as a replacement of expensive physical part deployment. Our research has leveraged the results from existing model validation work in the computational modeling community and extended their use in engineering design. Results are broadly disseminated throughout mechanical engineering, industrial engineering, simulation, and applied statistics communities. The strong collaborations between the research teams, industrial partners, and government agencies has ensured that the technology is transferred and the results are successfully implemented. The research has contributed to education in the areas of model-based simulation, modeling and optimization of engineering systems under uncertainty, statistical analysis, engineering design, and information technology as well as provide training to minority and women engineering students.

3. Research and Teaching Skills and Experience Provided

Graduate students supported under this grant had the opportunity to learn how to conduct collaborative research with researchers from a different research institution and with different background. Students were exposed to various issues related to probability and statistical analyses, engineering design, uncertainty modeling, etc. The project also provides the learning opportunity of presenting and publishing research results.

Regarding the education activities, the research results have directly benefited the teaching of ME495-Advanced Computational Methods for Engineering Design, a course taught at the graduate level. Model Validation is one of the several new topics added to this course.

The project has provided graduate students the opportunity of working on real world problems through collaborations with industry.

4. Outreach

In the course of this project, the research team has exchanged research ideas with many other research groups that have similar interests in the topic of model validation, which helps identify the research needs of the proposed project. Examples of these research groups include the Optimization and Uncertainty Estimation group at the Sandia National Laboratory, the Safety Engineering group in the Scientific Research Lab of Ford Motor Company, and the Global Performance Integration Group at General Motors. Research results have been presented at the Stochastic Modeling workshop at University of Notre Dame (March 24-26, 06), the Panel Session on "Transition of Non Deterministic Approaches from Academic and National Lab Research to Industrial Design and Decision-Making", at the SAE congress (April 6th, 06), the 2006, 2007, 2008 ASME Design Technical Conferences, the 7th World Congress on Structural and Multidisciplinary Optimization, the 2008 49th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, and the ASME 2008 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference.

5. Publications

[1] Wang, S., Chen, W., and Tsui, K, "Bayesian Validation of Computer Models," submitted to *Technometrics,* 2007.

[2] Chen, W., Tsui, K-L., Xiong, Y., and Wang, S., "Metric and a Bayesian Procedure for Validating Predictive Models in Engineering Design", *2006 ASME Design Technical Conference, Design Automation Conference*, September 10-13, Philadelphia, PA, 2006.

[3] Apley, D., Liu, J., and Chen, W., "Understanding the Effects of Model Uncertainty in Robust Design with Computer Experiments", *ASME Journal of Mechanical Design*, 128(4), 745-958, 2006.

[4] Xiong, Y., Chen, W., Apley, D., and Ding, X., "A Nonstationary Covariance Based Kriging Method for Metamodeling in Engineering Design", *11th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, Portsmouth, VA, September 6-8, 2006. In press, *International Journal for Numerical Methods in Engineering.*

[5] Xiong, Y., Chen, W., Tsui, K-L., "A New Variable Fidelity Optimization Framework Based on Model Fusion and Objective-Oriented Sequential Sampling", *ASME Design Engineering Technical Conference, Design Automation Conference*, September 4-7, Las Vegas, Nevada, 2007, accepted by ASME Journal of Mechanical Design.

[6] Xiong, Y., Chen, W., Apley, D., and Ding, X., "A Nonstationary Covariance-Based Kriging Method for Metamodeling in Engineering Design", *International Journal for Numerical Methods in Engineering,* 71(6), 733-756, August 2007.

[7] Lee, S. and Chen, W., "A Comparative Study of Uncertainty Propagation Methods for Black-Box Type Functions", accepted by *Structural and Multidisciplinary Optimization*, Jan. 2008.

[8] Yin, X. and Chen, W, "A Hierarchical Statistical Sensitivity Analysis Method for Complex Engineering Systems", *ASME Journal of Mechanical Design,* 130(7), 2008.

[9] Chen, W., Xiong, Y., Tsui, K-L., and Wang, S., "A Design-Driven Validation Approach using Bayesian Prediction Models", *ASME Journal of Mechanical Design*, 130(2), 2008.

[10] Xiong, Y., Chen, W, and Tsui, K., "A New Variable Fidelity Optimization Framework Based on Model Fusion and Objective-Oriented Sequential Sampling", accepted by *ASME Journal of Mechanical Design,* June 2008.

1. [11] Liu, X. Chen, W., Tsui, K-L., "Regression Modeling for Computer Model Validation with Functional Responses", Paper No. DETC2008-49662, *Proceedings of the ASME 2008 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference*, August 3-6, 2008, Brooklyn, New York.

2. [12] Xiong, Y., Chen, W., and Tsui, K-L, "A Better Understanding of Model Updating Strategies in Validating Engineering Models", *49th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Schaumburg, IL, Apr 7-10, 2008.

Acknowledgement: Support from CMMI #0522662 and #0522366 are gratefully acknowledged.

NSF GRANTS #0522662 and #0522366
NSF PROGRAM NAME: Engineering Design

Collaborative Research: Validating Predictive Models in Engineering Design

Wei Chen
Northwestern University

Kwok Tsui
Georgia Institute of Technology

Abstract: This collaborative research represents the joint efforts from two universities, Northwestern University (NU) and Georgia Tech (GT), for the development of a model validation approach that provides quantitative assessments of uncertainty in using predictive models in engineering design. During the first year of this project, the focus has been on implementing the following two research tasks: (1) Development of a Bayesian approach to assess the uncertainty in model prediction by combining data from both physical experiments and computer model; and (2) Development of a design-oriented model validation metric to guide designers for achieving high confidence of using predictive models in making design decision. The two teams at NU and GT have worked very closely on the above two tasks while each team takes the lead on one subject, i.e., GT for task (1) and NU for task (2). The research results have been documented in two joint publications.

1. A Bayesian approach to Uncertainty Quantification

Most research in validating computer models had focused on estimating *prediction bias* and improving accuracy of a computer model. Much less work had been done on characterizing *prediction uncertainty* and *prediction bias* under general situations. Using *x* to represent design variables and *y* stand for design performance, the relationship between the experimental observations $Y^e(\mathbf{x})$, the true behavior $Y^r(\mathbf{x})$, and the prediction generated by a computer model $Y^m(x)$ can often be generalized as follows:

$$Y^e(\mathbf{x}) = Y^r(\mathbf{x}) + \varepsilon(\mathbf{x}) = Y^m(\mathbf{x}) + \delta(\mathbf{x}) + \varepsilon(\mathbf{x}), \qquad (1)$$

where $\varepsilon(\mathbf{x})$ is the random variable representing the experimental error (relating to both experimental setup and measurement) that may depend on $\mathbf{x}$, and $\delta(\mathbf{x})$ is the error of the model , or called the prediction bias, i.e.,

$$\delta(\mathbf{x}) = Y^r(\mathbf{x}) - Y^m(\mathbf{x}), \qquad (2)$$

which captures the model inadequacy. The prediction bias $\delta(\mathbf{x})$ is more closely related to the assessment of model accuracy, while the prediction of the true model output $Y^r(\mathbf{x})$ is essential to assess the confidence of using model for decision making, measured by the probability that a design alternative will produce an outcome that is preferred to or indifferent to other alternatives.

We developed a Bayesian approach to provide uncertainty quantification of both $\delta(\mathbf{x})$ and $Y^r(\mathbf{x})$. Due to the lack of experimental data in most engineering design applications, we find that the Bayesian inferences may be preferred as they require fewer assumptions and are more flexible for applications, i.e., additional information can be incorporated through prior distributions. The proposed Bayesian procedure for model validation includes six major steps:

(1) Collect both physical and computer model data;
(2) Determine priors of Gaussian process parameters;
(3) Compute the posterior of computer model;
(4) Compute the posterior of prediction bias;
(5) Compute the prediction of the true behavior; and
(6) Evaluate predictive capability and design validity.

Mathematical details of uncertainty quantification of both $\delta(\mathbf{x})$ and $Y^r(\mathbf{x})$ can be found in Wang et al. (2006).

To illustrate this procedure, an engine piston design case is studied. The goal of the design is to optimize the skirt profile (SP, one of the key geometric parameters) of the engine piston to obtain the minimal piston slap noise. Over the considered design range, 34 hypothetical physical experiments (the circles in Figure 1) and 10 computer experiments (the triangles in Figure 1) are considered. Figures 1~ 3 show the results obtained from Steps (3)~(5), respectively.

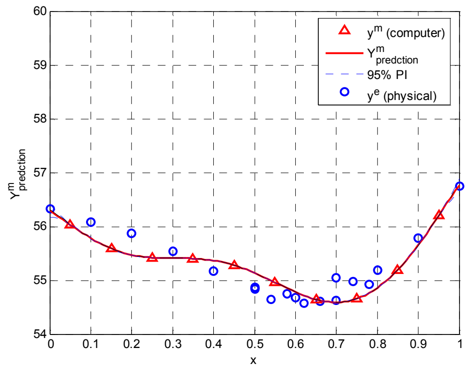

Figure 1. Posterior of computer model $\hat{Y}^m(\mathbf{x})$

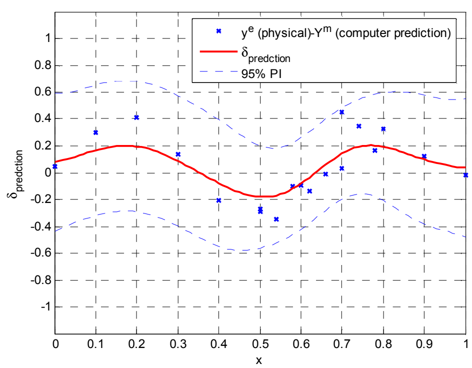

Figure 2. Posterior of prediction bias $\hat{\delta}(\mathbf{x})$

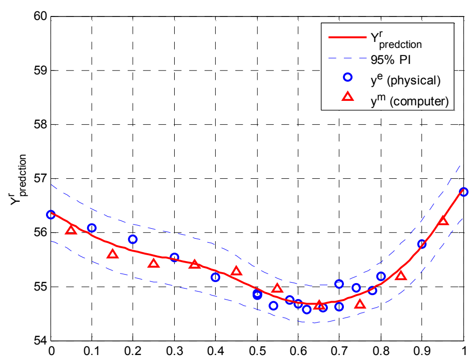

Figure 3. Prediction of the true behavior $\hat{Y}^r(\mathbf{x})$

2. Design Validation Metric

Different from the existing validation metrics that assess the predictive capability (accuracy) of a model, the design validation metrics M_D has been developed in our work to provide a probabilistic measure of whether the real outcome of a candidate design is better than other design choices. Such metric is developed to provide a direct measure of how reliable is the decision of choosing one design candidate versus the other design alternatives, therefore to provide the confidence associated with a design decision with the consideration of model uncertainty.

In addition, such metric provides useful guidance for validation activities. If large uncertainty exists in predicting design outcome, e.g., because the design sites are far from the tested region, the achieved M_D may be too low to meet the design validity requirements, forcing designers to add new experiments to reduce model uncertainty or to lower the validity requirement.

We have developed a design validation metric for choosing a particular design alternative by comparing it to a finite number (k) of other design alternatives [2]. The following three forms of the design validation metric are considered:

(1) The Multiplicative Metric: $M_D^{Multip}(\mathbf{x}_i)$

$$M_D^{Multip}(\mathbf{x}_i) = \prod_{j=1, j\neq i}^{k} P\{f^r(\mathbf{x}_i) < f^r(\mathbf{x}_j)\} \tag{3}$$

(2) The Average (Additive) Metric: $M_D^{Average}(x_i)$

$$M_D^{Average}(\mathbf{x}_i) = \frac{1}{k-1}\sum_{j=1, j\neq i}^{k} P\{f(\mathbf{x}_i) < f(\mathbf{x}_j)\} \tag{4}$$

(3) The Worst-Case Metric: $M_D^{Worstcase}(x_i)$

$$M_D^{Worstcase}(\mathbf{x}_i) = \min_{j=1,\ldots,k, j\neq i} P\{f(\mathbf{x}_i) < f(\mathbf{x}_j)\} \tag{5}$$

In Eqns. (3)~(5), $f^r(\mathbf{x}_i)$ stands for the prediction of the true design objective value, considering model uncertainty. Figure 4 illustrates the uncertainty quantification for a complex design objective in robust design that involves the assessment of both the mean and variance of performance. As shown in Figure 2, uncertainty in prediction can be reduced by adding more physical experiments. The results of design validation metric for each design alternative are provided in Table 1. For each alternative metric formulation used, $M_D(\mathbf{x}_4)$ for alternative $\mathbf{x}_4$ continues to be the largest one among the five alternatives, indicating the highest confidence of using the predictive model for selecting alternative $\mathbf{x}_4$.

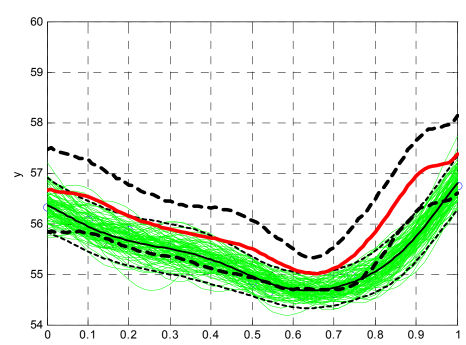

Figure 4. Uncertainty Quantification for Robust Design Objective, $f(\mathbf{x}) = \mu_{\hat{Y}^r}(\mathbf{x}) + k\sigma_{\hat{Y}^r}(\mathbf{x})$; Cluster of lines show realizations of $\hat{Y}^r(\mathbf{x})$; dark solid lines show

95% confidence interval $\mu_f(\mathbf{x}) \pm 2\sigma_f(\mathbf{x})$ for the robust design objective

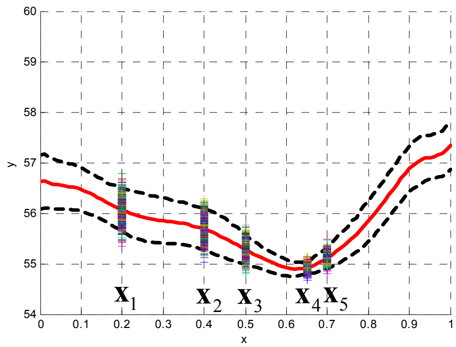

Figure 5. Uncertainty Quantification after Adding More Physical Experiments

Table 1 Design Validation Metric Values for Each Alternative

Design i	1	2	3	4	5
$M_D^{Multip}(\mathbf{x}_i)$	0.0000	0.0000	0.0155	**0.6553**	0.2244
$M_D^{Average}(\mathbf{x}_i)$	0.0093	0.2542	0.5551	**0.9069**	0.7744
$M_D^{Worstcase}(\mathbf{x}_i)$	0.0000	0.0030	0.0983	**0.7290**	0.2710

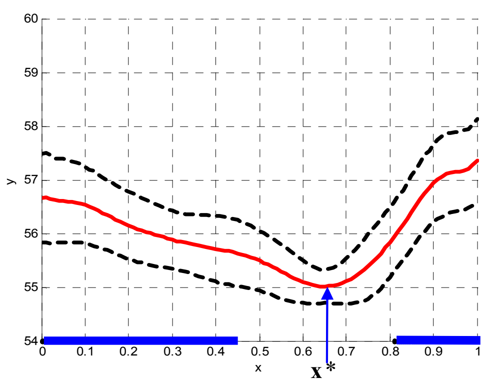

Figure 6. Identified inferior design region $\{\mathbf{x} \mid P\{f(\mathbf{x}^*) < f(\mathbf{x})\} > 0.95\}$ with respect to the optimal deign x_4

As shown in Figure 6, besides providing the confidence of accepting a design solution, we can also use the information of pair-wise comparison probability P_{ij} to rule out the inferior design region, therefore narrowing the sampling space in sequential samplings of physical and computer experiments.

3. Summary of Contributions

Compared to the existing work, our work results in a full Bayesian analysis model for predicting computer model bias and true model output, that are both accurate and economically sound. Our approach provides quantitative means to define and to assess model validity from the perspective of design decision making with the consideration of various sources of uncertainties. Our work offers a new and improved way of viewing model validation by relating its definition to a specific design choice. The proposed metric for assessing design validity provides probabilistic measurements with regard to the confidence of using a model for making a specific design choice; they can be used to overcome the limitations of many existing model validation approaches while providing direct estimate of the global impact of uncertainty sources on the confidence in a design decision.

Acknowledgement: Support from DMI- #0522662 and #0522366 are gratefully acknowledged.

References/Publications:

[1] Wang, S., Chen, W., and Tsui, K, "Bayesian Validation of Computer Models," working paper, to be submitted, 2006.

[2] Chen, W., Tsui, K-L., Xiong, Y., and Wang, S., "Metric and a Bayesian Procedure for Validating Predictive Models in Engineering Design", *2006 ASME Design Technical Conference, Design Automation Conference*, September 10-13, Philadelphia, PA, 2006.

Collaborative Research: Validating Predictive Models in Engineering Design

Wei Chen
Northwestern University

Kwok Tsui
Georgia Institute of Technology

Abstract

This collaborative research represents the joint efforts from two universities, Northwestern University (NU) and Georgia Tech (GT), for the development of a model validation approach that provides quantitative assessments of uncertainty in using predictive models in engineering design. Five major tasks have been implemented during the course of this project: (1) Development of a Bayesian approach to assess the uncertainty in model prediction by combining data from both physical experiments and computer model; (2) Development of a design-oriented model validation metric to guide designers for achieving high confidence of using predictive models in making design decision; (3) Development of objective-oriented sequential experiment design strategy; (4) Achieving a better Understanding of Model Updating Strategies in Validating Engineering Models; and (5) Modeling and validating computer models with functional responses. The two teams at NU and GT have worked very closely on all tasks. The research results have been documented in eighteen (18) journal/conference papers.

Research Findings

1. A Bayesian Approach to Model Uncertainty Quantification

Most research in validating computer models had focused on estimating *prediction bias* and improving accuracy of a computer model. Much less work had been done on characterizing *prediction uncertainty* and *prediction bias* under general situations. Using *x* to represent design variables and *y* stand for design performance, the relationship between the experimental observations $Y^e(\mathbf{x})$, the true behavior $Y^r(\mathbf{x})$, and the prediction generated by a computer model $Y^m(x)$ can often be generalized as follows:

$$Y^e(\mathbf{x}) = Y^r(\mathbf{x}) + \varepsilon(\mathbf{x}) = Y^m(\mathbf{x}) + \delta(\mathbf{x}) + \varepsilon(\mathbf{x}), \qquad (1)$$

where $\varepsilon(\mathbf{x})$ is the random variable representing the experimental error (relating to both experimental setup and measurement) that may depend on $\mathbf{x}$, and $\delta(\mathbf{x})$ is the error of the model , or the prediction bias, i.e.,

$$\delta(\mathbf{x}) = Y^r(\mathbf{x}) - Y^m(\mathbf{x}), \qquad (2)$$

which captures the model inadequacy. The prediction bias $\delta(\mathbf{x})$ is more closely related to the assessment of model accuracy, while the prediction of the true model output $Y^r(\mathbf{x})$ is essential to assess the confidence of using model for decision making, measured by the probability that a design alternative will produce an outcome that is preferred to or indifferent to other alternatives.

We developed a Bayesian approach to provide uncertainty quantification of both $\delta(\mathbf{x})$ and $Y^r(\mathbf{x})$. Due to the lack of experimental data in most engineering design applications, we find that the Bayesian inferences may be preferred as they require fewer assumptions and are more flexible for applications, i.e., additional information can be incorporated through prior distributions. The proposed Bayesian procedure for model validation includes six major steps:

(1) Collect both physical and computational data;
(2) Determine priors of Gaussian process parameters;
(3) Compute the posterior of computer model $Y^m(\mathbf{x})$;
(4) Compute the posterior of prediction bias $\delta(\mathbf{x})$;
(5) Compute the prediction of the true behavior $Y^r(\mathbf{x})$; and
(6) Evaluate predictive capability and design validity.

Mathematical details of uncertainty quantification of both $\delta(\mathbf{x})$ and $Y^r(\mathbf{x})$ can be found in Wang et al. [1]. We have provided a clear decomposition of the uncertainty of $Y^r(\mathbf{x})$. This decomposition explains why and how combining computer outputs and physical experiments can provide more accurate prediction than using only computer outputs or only physical experiments.

To illustrate the above six-step procedure, an engine piston design case is studied. The goal of the design is to optimize the skirt profile (SP, one of the key geometric parameters) of the engine piston to obtain the minimal piston slap noise (unit: dB). Over the considered design range, 6+3 physical experiments (the circles in Figure 1) and 9 computer experiments (the triangles in Figure 1, where the first 6 physical experiments are used) are considered. Figures 1~3 show the results obtained from Steps (3)~(5), respectively.

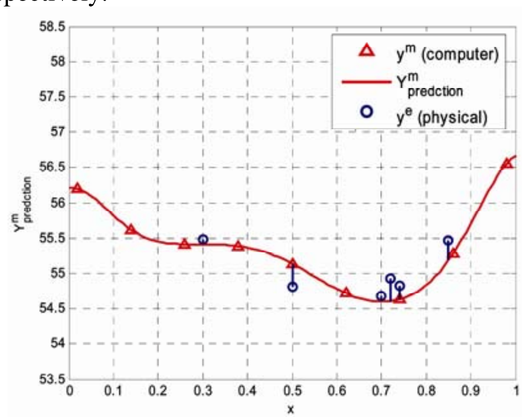

Figure 1. Posterior of computer model $\hat{Y}^m(\mathbf{x})$

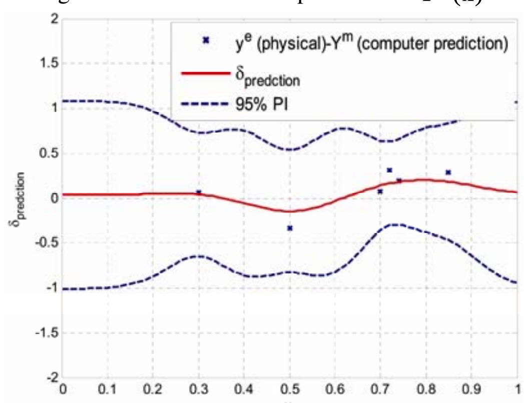

Figure 2. Posterior of prediction bias $\hat{\delta}(\mathbf{x})$

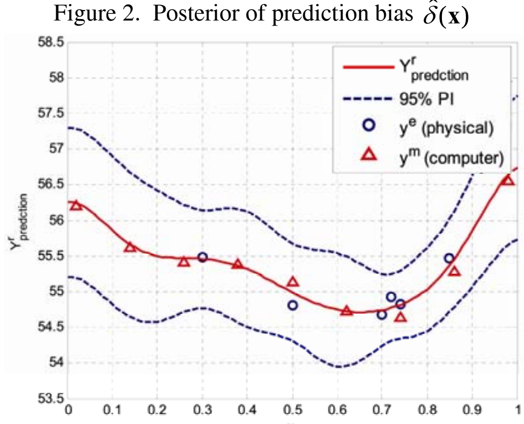

Figure 3. Prediction of the true behavior $\hat{Y}^r(\mathbf{x})$

2. Design Validation Metric

Different from the existing validation metrics that assess the predictive capability (accuracy) of a model, the design validation metrics M_D has been developed in our work to provide a probabilistic measure of whether the real outcome of a candidate design is better than other design choices. Such metric is developed to provide a direct measure of how reliable is the decision of choosing one design candidate versus the other design alternatives, therefore to provide the confidence associated with a design decision with the consideration of model uncertainty. In addition, such metric provides useful guidance for validation activities. If large uncertainty exists in predicting design outcome, e.g., because the design sites are far from the tested region, the achieved M_D may be too low to meet the design validity requirements, forcing designers to add new experiments to reduce model uncertainty or to lower the validity requirement.

We have developed a design validation metric for choosing a particular design alternative by comparing it to a finite number of other design alternatives [2]. The following three forms of the design validation metric are considered:

(1) The Multiplicative Metric:

$$M_D^M(\mathbf{x}^*) = \left\{ \prod_{\mathbf{x}_i \in \Omega_d, \mathbf{x}_i \notin X^0} P\left\{ f(\mathbf{x}^*) < f(\mathbf{x}_i) \right\} \right\}^{1/K} \tag{3}$$

(2) The Average (Additive) Metric:

$$M_D^A(\mathbf{x}^*) = \frac{1}{K} \sum_{\mathbf{x}_i \in \Omega_d, \mathbf{x}_i \notin X^0} P\left\{ f(\mathbf{x}^*) < f(\mathbf{x}_i) \right\} \tag{4}$$

(3) The Worst-Case Metric:

$$M_D^W(\mathbf{x}^*) = \min_{\mathbf{x}_i \in \Omega_d, \mathbf{x}_i \notin X^0} P\left\{ f(\mathbf{x}^*) < f(\mathbf{x}_i) \right\} \tag{5}$$

The probability $P\left\{ f(\mathbf{x}^*) < f(\mathbf{x}_i) \right\}$ is evaluated for comparing $\mathbf{x}$* against individually each other design alternatives, where $\mathbf{x}_i$ (i=1,2,..,K) belongs to the set of feasible design alternatives Ω_d , excluding those in a indifferentiable region X^0. The concept of indifferentiable region X^0 is introduced to consider the fact that, with the consideration of model uncertainty, distinguishing designs with identical mean values might not be possible.

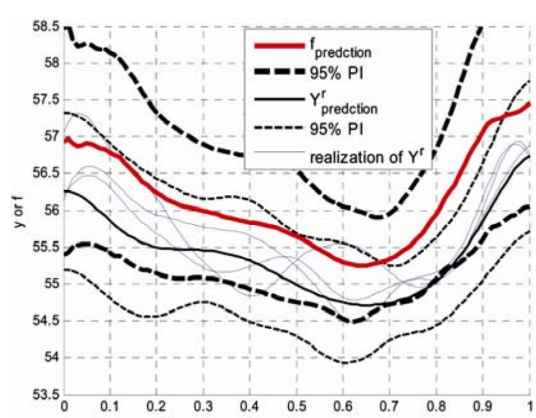

Figure 4. Prediction of $\hat{f}(\mathbf{x})$ (dB) and 95% confidence interval (w/ 6 physical experiments)

Figure 4 illustrates the uncertainty quantification for a complex design objective $f(\mathbf{x}) = \mu_{\hat{Y}^r}(\mathbf{x}) + k\sigma_{\hat{Y}^r}(\mathbf{x})$ in a typical robust design that involves the assessment of both the mean and variance of performance. Thin lines show three examples of realizations of $\hat{Y}^r(\mathbf{x})$; dark dashed lines show 95% confidence interval $\mu_f(\mathbf{x}) \pm 2\sigma_f(\mathbf{x})$ for $f(\mathbf{x})$.

We apply the decision validation metrics M_D proposed to the robust engine piston design. Two design scenarios, namely, discrete design alternatives and continuous design space, are considered separately.

Scenario 1. Discrete Design Alternatives

Figures 5 shows the mean and 95% prediction interval of $Z(\mathbf{x}_i) = f(\mathbf{x}_i) - f(\mathbf{x}^*)$, at five candidate points $\mathbf{x}_i$={0.2, 0.4, 0.5, 0.65, 0.7}. Note both the mean and variance of $Z(\mathbf{x}_i)$ reduces to zero at **x***=$\mathbf{x}_4$.

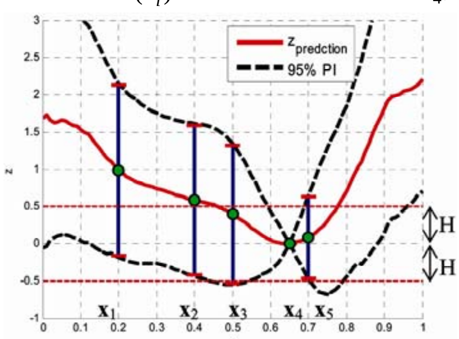

Figure 5. Mean and 95% CI of $Z(\mathbf{x}_i)$ (dB) at five design candidates (w/ 6 physical experiments)

Tables 1 and 2 show the results of decision validation metrics M_D with 6 and $6+3$ physical experiments. Comparing Tables 1 and 2, we find that values of M_D^M and M_D^W are higher in Table 2 than Table 1, reflecting the improved confidence of claiming $\mathbf{x}_4$ as the optimal design. In Table 2, with the reduced uncertainty in $Z(\mathbf{x}_i)$ by the inclusion of 3 additional experiments, it is found $P[|Z(\mathbf{x}_5)| < H]$ = 0.9799, indicating that $\mathbf{x}_5$ should be claimed indifferentiable to **x***=$\mathbf{x}_4$, and excluded for calculating M_D. Because $\mathbf{x}_5$ is considered as indifferentiable to **x***=$\mathbf{x}_4$, $\mathbf{x}_3$ becomes the most competing design to **x*** under the worst-case metric.

Among the three types of validation metrics, we found that the Worst-Case Metric M_D^W holds the most straightforward meaning to reflect the confidence of claiming an optimal design **x***, because it only concerns the most competitive design (2nd best design $\mathbf{x}_i$) to **x***, outside the indifferentiable region X^0. Besides, M_D^W is the easiest to implement for problems with a continuous design space. Optimization could be used to locate the worst-case point $\mathbf{x}_i$, by taking $P[Z(\mathbf{x}) > 0]$ as the objective to minimize and treating $P[|Z(\mathbf{x})| < H] < c$ (i.e., $\mathbf{x} \notin X^0$) and $\mathbf{x} \in \Omega_d$ as the constraints.

Table 1. Calculation of the decision validation metrics (H=0.5, c=95%, 6 physical experiments)

$\mathbf{x}_i$	$\mathbf{x}_1$	$\mathbf{x}_2$	$\mathbf{x}_3$	$\mathbf{x}_4$	$\mathbf{x}_5$
$P[\lvert Z(\mathbf{x}_i)\rvert < H]$	0.1968	0.4152	0.5568		0.9114
$\mathbf{x}_i \in X^0$?	No	No	No		No
$P[Z(\mathbf{x}_i) > 0]$	0.9544	0.8742	0.7981		0.6204
M_D^M	0.8017				
M_D^A	0.8118				
M_D^W	0.6204 (worst-case point: $\mathbf{x}_5$)				

Table 2. Calculation of the decision validation metrics (H=0.5, c=95%, 6+3 physical experiments)

$\mathbf{x}_i$	$\mathbf{x}_1$	$\mathbf{x}_2$	$\mathbf{x}_3$	$\mathbf{x}_4$	$\mathbf{x}_5$
$P[\lvert Z(\mathbf{x}_i)\rvert < H]$	0.1910	0.3790	0.6111		0.9799
$\mathbf{x}_i \in X^0$?	No	No	No		Yes
$P[Z(\mathbf{x}_i) > 0]$	0.9616	0.9034	0.7826		
M_D^M	0.9080				
M_D^A	0.6619				
M_D^W	0.7826 (worst-case point: $\mathbf{x}_3$)				

Scenario 2. Continuous Design Space

Calculating M_D in a continuous design space is more challenging than in a discrete design space. Figures 6(a) and 6(b) show respectively the mean and 95% confidence interval of $Z(\mathbf{x})$ for two experimental sizes, 6 and 6+3, where the solid bold portions indicate the indifferentiable region X^0 in the small neighborhood of **x*** (identified by minimizing $\hat{f}(\mathbf{x}_i)$).

Table 3 provides the calculated M_D values. The increased M_D values with more physical experiments reflect the improved confidence of claiming **x*** as the optimal design after the resolution of $f(\mathbf{x})$ is improved. Table 3 also provides the calculated M_D values when the tolerance H is set at a higher value, 0.9 (dB). M_D^M and M_D^W increase as a result of the less strict tolerance: when a larger tolerance H is specified by a designer, it implies that lower resolution of the model is demanded.

Table 3. Calculation of decision validation metrics at different tolerance H and experiment size (w/ c=95%)

	H=0.5(dB)		H=0.9(dB)	
Phy. exp. #	6	6+3	6	6+3
M_D^M	0.8953	0.9203	0.9281	0.9535
M_D^A	0.8122	0.8161	0.7516	0.7457
M_D^W	0.5947	0.6289	0.6706	0.7341

WC point	**x**=0.69	**x**=0.59	**x**=0.73	**x**=0.53

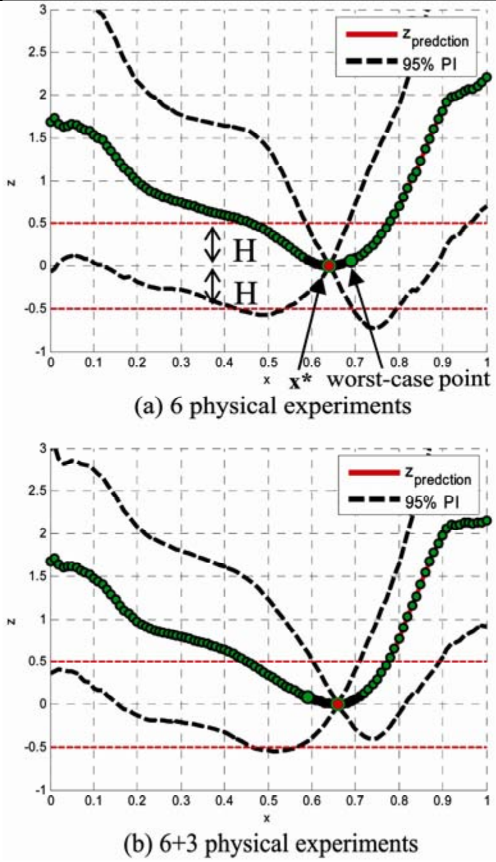

Figure 6. Mean and 95% confidence interval of $Z(\mathbf{x})$ (dB) (H=0.5, c=95%)

3. Sequential Experiment Design

A sequential sampling strategy is first developed for computer experiments in variable fidelity optimization [3]. We applied the Bayesian approach developed in the model validation research to model fusion for integrating high fidelity (HF) and low fidelity (LF) models into the predicative surrogate model, over which design optimizations are performed. The developed sequential sampling strategy is intended to overcome existing sequential sampling methods. A periodical switching criteria (PSC) strategy is proposed, which is depicted in Figure 7. To locate the next computer experiment $\boldsymbol{x}_{N+1}$, the Statistical Lower Bounding (SLB) criterion is used with the form

$$SLB(\mathbf{x}_{N+1}) \equiv \mu_{y^s}(\mathbf{x}_{N+1}) - k\sigma_{y^s}(\mathbf{x}_{N+1}) \,. \quad (6)$$

Compared with the conventional Expected Improvement (EI) method, The SLB criterion is easier to control by adjusting k, a parameter interpreted as a weight placed for reducing interpolation uncertainty against enhancing the local accuracy around the current optimum.

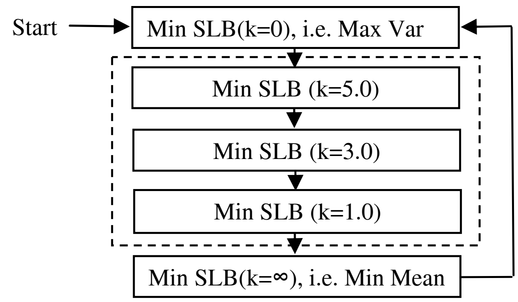

Figure 7. The proposed periodical switching criteria (PSC) strategy for sequential experiment design

The developed methodology of sequential sampling in variable fidelity optimization can be extended to model validation, by treating LF simulation as computational experiments, HF simulations as physical experiments with diminished experiment error. In our proposed design-driven model validation framework (Figure 8), single/multiple physical experiment(s) could be identified by PSC strategy and added sequentially. The sequential validation process is repeated until the validity requirement is satisfied, which is determined by examining if the achieved decision validation metric M_D has reached a prespecified threshold P_{th}. With the proposed validation metric, we are able to enhance the predictive capability of a computer model for the purpose of design decision making. One important feature of the proposed sequential experiment design is that, by treating model uncertainty separately from design variable/parameter uncertainty, we are able to effectively design the physical experiments, to sequentially eliminate the model uncertainty.

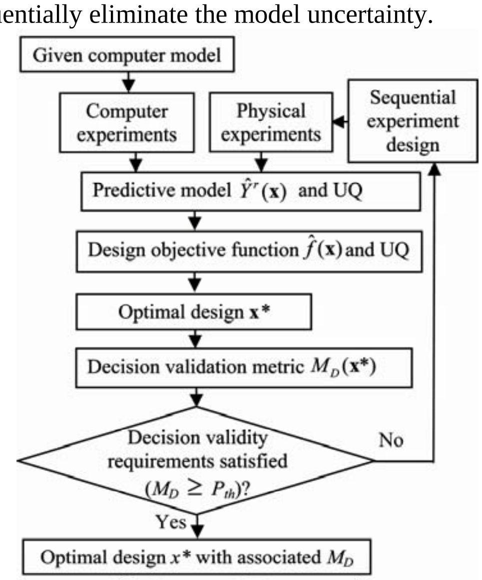

Figure 8. The proposed design-driven validation framework with sequential experiment design

4. A better Understanding of Model Updating Strategies in Validating Engineering Models

In this task, we examine various model updating strategies as an integral part of the model validation process [13]. The existing model updating strategies differ in their formulations, the solution method used, and the physical interpretations. The two most widely used categories of formulations include bias-correction and calibration.

Through our examination, we found there are several limitations when applying the traditional Bayesian calibration approaches to update a computer model using either bias-correction, calibration, or a combination of both. Besides the numerical difficulty in implementation, one major limitation of the Bayesian approach is that the calibration parameters are treated as uncertain due to lack of knowledge, not accounting for sources of variability in a validation process. Besides, the choice of the prior distribution in Bayesian analysis is often arbitrary.

As an alternative approach to the traditional Bayesian approach, we examine in this task a new model updating strategy, in which a computer model is updated to better interpret the observed *dispersion* of experimental data. The Maximum Likelihood Estimation (MLE) method is used to estimate the model updating parameters as shown in Figure 9. Unlike the traditional Bayesian approach which accounts for the experimental uncertainty by a single error term, the MLE based model updating approach accounts for experimental uncertainty through a subset of model updating parameters. In contrast to the Bayesian approach, the MLE based approach does not rely on the prior distributions of calibration parameters, instead, it seeks optimal distribution parameters underlying model updating parameters through maximizing the likelihood function based on the physical experiment data (Figure 10).

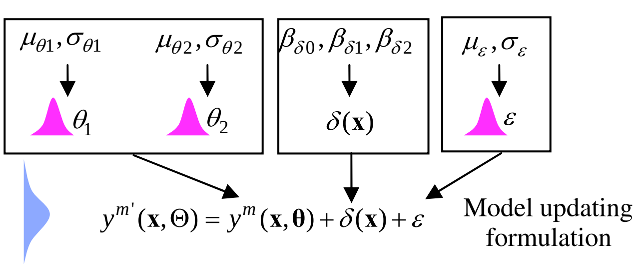

Figure 9. Model updating parameters Θ in formulation $y^{m'}(\mathbf{x},\Theta)$

Through the thermal challenge example, we demonstrate that model updating can be treated as an integral part of a model validation process which improves a model based on the physical observations gathered. We illustrate that without running into numerical complexity, the model updating method we proposed is easier to implement and interpret compared to the existing Bayesian methods.

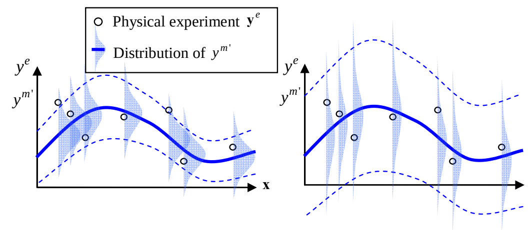

Figure 10. Likelihood value indicates the agreement between the output distribution of the updated model and the dispersion of physical experiments

Using the newly developed u-pooling method by Ferson et al, we show that the metric can be applied to both the original and the updated models to assess the accuracy and predictive capability of different model updating formulations. Through in-sample and out-sample tests (Figure 11) based on different data sets, we find that the proposed model updating approach improves the agreement between the model and the physical experiment data. However, when applying the updated model at a region that is far from the domain of data used for model updating, the extrapolation capability of the updated model is not guaranteed.

By comparing our approach to the existing works on the thermal challenge problem, we observe the differences of various methods in utilizing available data, the model updating formulations adopted, and the solution method employed. Even though our method is different, we find the conclusion we reach on device failure probability is identical to other methods in literature. As for which model updating formulation is the most appropriate, unless it can be specified based on the pre-existing knowledge, we think it is problem dependent and should be selected by exercising the model validation metrics as demonstrated.

While model updating is shown to be useful for improving the accuracy of a model, as the process is fully data-driven, we believe the method should be used with caution when used for extrapolation.

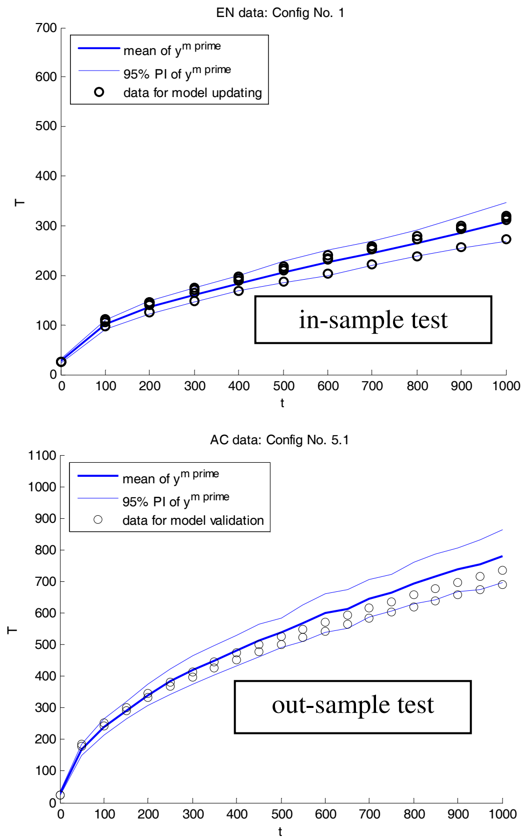

Figure 11. Example of In-Sample and Out-Sample Test

Due to the nature of the MLE method, the effectiveness and accuracy of the MLE based model updating approach could be downgraded when data amount is extremely small. In our test with the 'low level' data sufficiency for the thermal challenge problem, it is found that the bandwidth of the prediction uncertainty could be degenerated to fairly small values, unable to reflect the condition of lack of data. To mitigate this problem, prior knowledge may be used to specify more conservative bounds of model updating parameters to prevent them from running into 'absurd' values. Another potential weakness of the MLE based model updating approach might be associated with the numerical instability during the optimization of the likelihood function, especially when a complex model updating formulation that involves many parameters is considered. To mitigate this issue, sensitivity analysis could be performed prior to optimization, by leaving out parameters that are insensitive to model output and the likelihood function.

5. Modeling and validating computer models with functional responses.

Statistical analysis of functional responses based on functional data from both computer and physical experiments has gained increasing attention due to the dynamic nature of many engineering systems. However, the complexity and huge amount of functional data bring many difficulties to apply traditional or existing methodologies. The objective of the present study is twofold: (1) prediction of functional responses based on functional data and (2) prediction of bias function for validation of a computer model that predicts functional responses.

A single step functional regression modeling approach is developed under this task to analyze functional outputs of physical and computer experiments [11]. Traditional methods for modeling functional data generally involve two steps. Models are first fit at each individual setting of the input to reduce the dimensionality of the functional data. Then the estimated parameters of the models are treated as new responses, which are further modeled for prediction. Alternatively, pointwise models are first constructed at each time point and then functional curves are fit to the parameter estimates obtained from the fitted models. We propose a single model to relate the functional response to both the input and the time variables. To overcome the high correlation between the shift (ending) response and the shift time, a sequential procedure is proposed to model the shift time as a function of the inputs and the shift response. We find the proposed model may be easier to interpret and implement for certain applications. Through a comparison with the existing Gaussian process modeling approach using a real industrial example provided by General Motor, we demonstrate that the proposed method yields sufficient accuracy, performs efficiently and achieves satisfactory accuracy in global prediction.

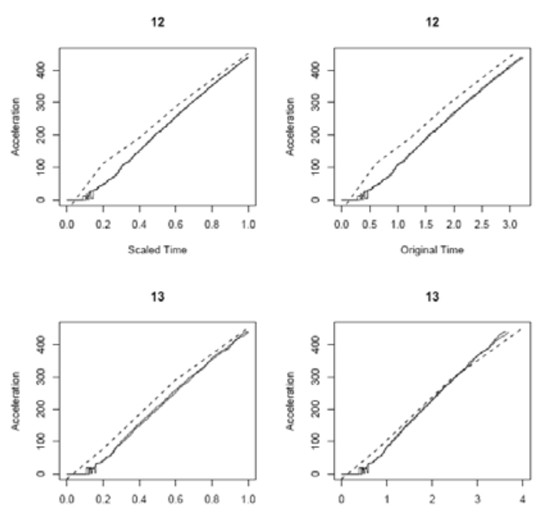

Figure 12: Predictions of two untested conditions: Observed (solid line), Predicted (dashed line)

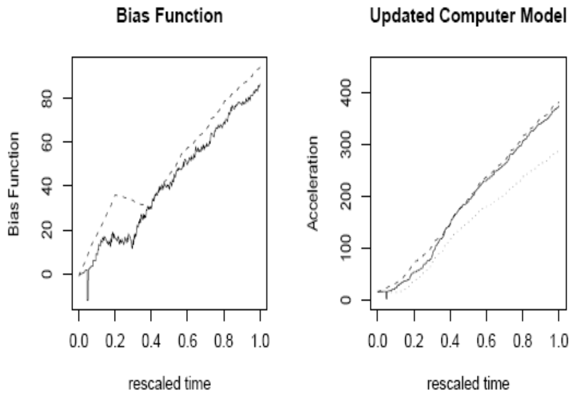

Figure 13: Bias Function and Updated Computer Outputs (Prediction: Dashed; Physical (or real bias function): Solid; Computer: Dotted)

Summary of Contributions

1. Intellectual Contributions

Compared to the existing work, our work focuses on a Bayesian model for predicting computer model bias and true model output, that are accurate, flexible and economically sound. In engineering applications where it is too expensive to obtain experimental data, the Bayesian inference approach offers much flexibility as additional design knowledge and information can be easily incorporated through prior distributions. With the Bayesian approach, uncertainty in prediction related to the lack of experiment data can be captured by the magnitude of uncertainty of the bias function, which offers rigorous and flexible methods for quantifying the model uncertainty in an intended design domain that may interpolate as well as extrapolate from a validation domain. Since we have developed the analytical results in implementing the Bayesian approach, the Bayesian approach we proposed can be economically implemented in multidimensional problems.

Based on the Bayesian approach we proposed, our research is the first work that provides theoretical discussion on the significance of combining computer outputs and physical, which can improve the prediction of the real system output over using only computer outputs or only physical observations.

Our research is one of the pioneering works that provide quantitative means to define and to assess model validity from the perspective of design decision making with the consideration of various sources of uncertainties. It offers a new and improved way of viewing model validation by relating its definition to a specific design choice. The proposed metric for assessing design validity provide probabilistic measurements with regard to the confidence of using a model for making a specific design choice; they can be used to overcome the limitations of many existing model validation approaches while providing direct estimate of the global impact of uncertainty sources on the confidence in a design decision.

Our research also clarified the role that model validation should play in decision making under uncertainty and developed strategies for making tradeoffs based on both product design and model validation. Unlike most of the existing model validation works that focus on the assessment of model accuracy, model validation in our research is viewed as a process to improve designer’s confidence in making a design choice using the improved predictive model, which is the augmented model that includes both the original computer model as well as the estimate of the bias function. The research directly addressed the needs of Engineering Design programs that seek improvement on validation of models, increased emphasis on treatment of uncertainty, and improvement on computational tools needed to implement the theory.

Our research proposed a new and effective strategy of sequential experiment design in variable fidelity optimization, which is immediately extendable to the model validation. Using decision validation metrics for assessing the confidence of the optimum design, we are able to enhance the predictive capability of a computer model for the purpose of design decision making. By treating model uncertainty separately from design variable/parameter uncertainty, we are able to effectively design the physical experiments, to sequentially eliminate the model uncertainty.

There is a growing recognition that a model needs to be updated to better reflect the physical experiment observations that are collected in model validation. Our research provides a better understanding of the various model updating strategies, which utilize mathematical means to update a computer model based on both physical and computer observations. The Maximum Likelihood Estimation (MLE) method proposed provides a better interpretation of the observed dispersion of experimental data. Uncertainty in model prediction is quantified to account for various sources of uncertainty in a validation process. Since our approach is applied to the widely used benchmark thermal challenge problem, other researchers who are interested in this topic can further compare our results with those from their studies. The research provides more insights into the benefits and limitations of using the MLE method versus the Bayesian approach. Insights into various model updating strategies are also obtained through this study and can serve as the guideline in engineering practice.

Following the classical nonparametric regression framework, our proposed method for modeling and validating functional response uses a single step procedure which is shown to be easily implemented and computationally efficient.

3. Significance (Impact)

Our research has offered a generic model validation approach that can be applied to many domestic and military applications for making reliable decisions when using predictive models as a replacement of expensive physical part deployment. Our research has leveraged the results from existing model validation work in the computational modeling community and extended their use in engineering design. Results are broadly disseminated throughout mechanical engineering, industrial engineering, simulation, and applied statistics communities. The strong collaborations between the research teams, industrial partners, and government agencies has ensured that the technology is transferred and the results are successfully implemented. The research has contributed to education in the areas of model-based simulation, modeling and optimization of engineering systems under uncertainty, statistical analysis, engineering design, and information technology as well as provide training to minority and women engineering students.

3. Research and Teaching Skills and Experience Provided

Graduate students supported under this grant had the opportunity to learn how to conduct collaborative research with researchers from a different research institution and with different background. Students were exposed to various issues related to probability and statistical analyses, engineering design, uncertainty modeling, etc. The project also provides the learning opportunity of presenting and publishing research results.

Regarding the education activities, the research results have directly benefited the teaching of ME495-Advanced Computational Methods for Engineering Design, a course taught at the graduate level at Northwestern University. Model Validation is one of the several new topics added to this course. Based on the research results, Professor Chen also developed the teaching materials on model validation and uncertainty quantification for a newly founded interdisciplinary doctoral cluster on 'Predictive Science & Engineering Design (PSED) at Northwestern University. As the topic of validating engineering models based on both computer simulations and physical experiments has become the core issue of using science-based predictive models across multiple engineering and science fields, this research has helped to establish the theoretical foundation for teaching the principles of model validation and uncertainty quantification across multiple fields.

The project has provided graduate students the opportunity of working on real world problems through collaborations with industry.

4. Outreach

In the course of this project, the research team has exchanged research ideas with many other research groups that have similar interests in the topic of model validation, which helps identify the research needs of the proposed project. Examples of these research groups include the Optimization and Uncertainty Estimation group at the Sandia National Laboratory, the design methodology group at Ecole Central Paris, the Safety Engineering group in the Scientific Research Lab of Ford Motor Company, and the Global Performance Integration Group at General Motors. Research results have been presented at the Stochastic Modeling workshop at University of Notre Dame (March 24-26, 06), the Panel Session on 'Transition of Non Deterministic Approaches from Academic and National Lab Research to Industrial Design and Decision-Making', at the SAE congress (April 6th, 06), the 2006, 2007, 2008, 2009 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference., the 7th World Congress on Structural and Multidisciplinary Optimization, the 2008 49th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, the 2006, 2007, 2008 INFORMS Annual Conference, and the First, Second, and Third Pre-Conference Workshop of Data Mining. Professors Chen and Tsui have also delivered invited talks on the research subject at a number of university seminars (e.g., Purdue University, University of Florida, University of Texas-Arlington, Virginia Tech, Chinese University of Hong Kong, City University of Hong Kong, Hong Kong University of Science and Technology, Shanghai Jiao Tong University, Tong Ji University, University of Electronic Science and Technology of China) and industry visits (e.g., Ford Motor, General Motors, Boeing, and General Electric).

5. Publications

[1] Wang, S., Tsui, K-L., and Chen, W., "Bayesian Validation of Computer Models", *Technometrics*, to appear in November 2009.

[2] Chen, W., Tsui, K-L., Xiong, Y., and Wang, S., "Metric and a Bayesian Procedure for Validating Predictive Models in Engineering Design", *2006 ASME Design Technical Conference, Design Automation Conference*, September 10-13, Philadelphia, PA, 2006.

[3] Apley, D., Liu, J., and Chen, W., "Understanding the Effects of Model Uncertainty in Robust Design with Computer Experiments", *ASME Journal of Mechanical Design*, 128(4), 745-958, 2006.

[4] Xiong, Y., Chen, W., Apley, D., and Ding, X., "A Nonstationary Covariance Based Kriging Method for

Metamodeling in Engineering Design", *11th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, Portsmouth, VA, September 6-8, 2006.

[5] Xiong, Y., Chen, W., Tsui, K-L., "A New Variable Fidelity Optimization Framework Based on Model Fusion and Objective-Oriented Sequential Sampling", *ASME Design Engineering Technical Conference, Design Automation Conference*, September 4-7, Las Vegas, Nevada, 2007, accepted by ASME Journal of Mechanical Design.

[6] Xiong, Y., Chen, W., Apley, D., and Ding, X., "A Nonstationary Covariance-Based Kriging Method for Metamodeling in Engineering Design", *International Journal for Numerical Methods in Engineering,* 71(6), 733-756, August 2007.

[7] Lee, S. and Chen, W., "A Comparative Study of Uncertainty Propagation Methods for Black-Box Type Functions", accepted by *Structural and Multidisciplinary Optimization*, Jan. 2008.

[8] Yin, X. and Chen, W, "A Hierarchical Statistical Sensitivity Analysis Method for Complex Engineering Systems", *ASME Journal of Mechanical Design,* 130(7), 2008.

[9] Chen, W., Xiong, Y., Tsui, K-L., and Wang, S., "A Design-Driven Validation Approach using Bayesian Prediction Models", *ASME Journal of Mechanical Design*, 130(2), 2008.

[10] Xiong, Y., Chen, W, and Tsui, K., "A New Variable Fidelity Optimization Framework Based on Model Fusion and Objective-Oriented Sequential Sampling", *ASME Journal of Mechanical Design*, 130(11), 111401 (9 pages), 2008.

[11] Liu, X. Chen, W., Tsui, K-L., "Regression Modeling for Computer Model Validation with Functional Responses", Paper No. DETC2008-49662, *Proceedings of the ASME 2008 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference*, August 3-6, 2008, Brooklyn, New York.

[12] Xiong, Y., Chen, W., and Tsui, K-L, "A Better Understanding of Model Updating Strategies in Validating Engineering Models", *49th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference,* Schaumburg, IL, Apr 7-10, 2008.

[13] Xiong, Y., Chen, W., Tsui, K-L., and Apley, D., "A Better Understanding of Model Updating Strategies in Validating Engineering Models", *Journal of Computer Methods in Applied Mechanics and Engineering*, 198 (15-16), pp. 1327-1337, March 2009.

[14] Chen, S., Chen, W., and Lee, S., "Level Set Based Robust Shape and Topology Optimization under Random Field Uncertainties", accepted by *Structural and Multidisciplinary Optimization*, October 2009.

[15] Xiong, F., Greene, S., Chen, W., and Yang, S., "A New Sparse Grid Based Method for Uncertainty Propagation", in press *Structural and Multidisciplinary Optimization,* August 2009.

[16] Xiong, F., Xiong, Y., Chen, W., and Yang, S., "Optimizing Latin Hypercube Design for Sequential Sampling of Computer Experiments", *Engineering Optimization*, 41(8), 793-810, 2009.

[17] Chen, S, Xiong, Y., and Chen, W., "A Multiresponse and Multistage Metamodeling Approach for Design Optimization", *AIAA Journal*, 47(1), January 2009.

[18] Lee, S. and Chen, W. "A Comparative Study of Uncertainty Propagation Methods for Black-Box Type Functions", *Structural and Multidisciplinary Optimization*, 37(3), 239-253. 2008.

Acknowledgement: Support from CMMI #0522662 and #0522366 are gratefully acknowledged.