Reduction of Epistemic Uncertainty in Multifidelity Simulation-Based Multidisciplinary Design

Dr. Wei Chen
Wilson-Cook Professor in Engineering Design
Dr. Zhen Jiang (Ford Research and Innovation Center)
Dr. Shishi Chen (Beijing Institute of Technology)
Professor Daniel W. Apley (Industrial Engineering)
Northwestern University, Evanston, IL, USA
Simulation-Based Design under Uncertainty

**SIMULATION-BASED DESIGN**

- An information-seeking and learning process

- **Aleatory uncertainty**
  - Due to natural/physical randomness; irreducible

- **Epistemic uncertainty**
  - Due to lack of data and/or knowledge; reducible

**SOURCES OF UNCERTAINTY THAT AFFECT MODEL PREDICTION**

- **Epistemic Uncertainty**
  - Model bias
  - Parameter uncertainty
    - Due to naturally fixed but unknown model parameters
  - Interpolation uncertainty
    - Due to lack of data
  - Numerical uncertainty
    - Due to numerical implementations of a model

- **Aleatory Uncertainty**
  - Input variability
    - Operating conditions; manufacturing ...
  - Experimental variability

**DESIGN UNDER UNCERTAINTY**

To achieve a design that is insensitive to uncertainties

Multidisciplinary Design Optimization (MDO)

- Requires analyses in multiple disciplines
  Involves multiple subsystems and/or components

- Fusion SE 2014 image from Ford Motor Co
- FEA model images provided by Dr. Lei Shi, Shanghai Jiao Tong University
- Control system image from StabiliTrak
Multidisciplinary Design Optimization (MDO)

- Requires analyses in multiple disciplines
  Involves multiple subsystems and/or components

  **CHALLENGE #1**
  Coupling in analysis and UQ

  **CHALLENGE #2**
  Dynamic decision making in resource allocation

- Interdisciplinary couplings
  - Feed-forward Coupling
  - Feedback Coupling
Multiple Models with Different Levels of Fidelity

- **CHALLENGE #3**
  Heterogenous information from different sources (multifidelity simulations and experiments)

- **“High-fidelity”** physics-based CAE model
- **“Intermediate-fidelity”** physics-based CAE model
- **“Low-fidelity”** simplified handbook equations

\[
\begin{align*}
2m_c v_i^2 &= \frac{1}{2} (2m_c + m_s) v_j^2 + E_{structural} \\
2m_c v_i^2 &= \frac{1}{2} (2m_c + m_s) \left( \frac{2m_c v_j}{2m_c + m_s} \right)^2 + E_{structural} \\
E_{structural} &= 2m_c v_i^2 - 2m_c v_j^2 \\
E_{structural} &= 2m_c v_i^2 \left( 1 - \frac{m_c}{2m_c + m_s} \right)
\end{align*}
\]
Model-Fusion for Combining Heterogeneous Information
- Both hierarchical and nonhierarchical rankings of fidelity

Managing Couplings and Information Complexity
- Multidisciplinary statistical sensitivity analysis (MSSA)
- Multidisciplinary uncertainty analysis (MUA)

Resource Allocation for Reducing Epistemic Uncertainty in MDO
- How to design paths of information seeking actions
- Decision making meta-optimization problem
Any pair of random variables, \( Y(\mathbf{x}) \) and \( Y(\mathbf{x}') \), is spatially correlated.

Example: Gaussian Process

\[
Y(\mathbf{x}) \sim GP \left( m(\mathbf{x}), V(\mathbf{x}, \mathbf{x}') \right)
\]

\[
m(\mathbf{x}) = \mathbf{h}(\mathbf{x})^T \beta, \quad V(\mathbf{x}, \mathbf{x}') = \sigma^2 \exp \left\{ -\sum_k \omega_k (x_k - x'_k)^2 \right\}
\]

Introduction to Spatial Random Process Based Model Uncertainty Quantification

95% prediction interval quantifies uncertainty in the response prediction.

Observed data

Predicted response = posterior mean

Low-fidelity model

Prediction mean

Tests / High-fidelity model

--- 95% PI (prediction interval)
TOPIC 1

Model Fusion for Combining Heterogeneous Information

The fidelity levels of the simulation models can be clearly identified and then preliminarily ordered for a hierarchical model updating.


- Apply low-fidelity information to construct the approximation space for a high-fidelity surrogate and then compute a high-fidelity reconstruction for model prediction.
- Using stochastic allocation with generalized polynomial chaos approach.


- Assume the higher-fidelity model to be approximated by its next lower-fidelity model with a discrepancy, and then construct a multi-model sequential updating framework.
- Apply spatial random process (SRP) to surrogate the responses from different models.

Common Assumption

- The fidelity levels of the simulation models can be clearly identified and then preliminarily ordered for a hierarchical model updating.
Goal of this work: Develop model fusion techniques with uncertainty quantification for combining information from multiple models without a clear ranking of fidelity.
Three Spatial Random Process (SRP) based Approaches

Approach 1: Weighted Sum

\[ y^t(x) = y^e(x) - \varepsilon = \sum_i \rho^{(i)} m^{(i)}(x) + \delta(x) \]

Assumption: Independency between simulations and the discrepancy function

\[ \text{Cov} \left( m^{(i)}(x), \delta(x) \right) = 0, \quad \forall i \]

\( y^t(x) \): True response
\( y^e(x) \): Experimental response
\( m^{(i)}(x) \): \( i \text{th} \) simulation model
\( \delta(x) \): Discrepancy function
\( \varepsilon \): Experimental error

Approach 2: Each Model Individually Corrected

\[ y^t(x) = y^e(x) - \varepsilon = m^{(i)}(x) + \delta^{(i)}(x) \]

Assumption: Independency between the discrepancy function and the true response

\[ \text{Cov} \left( y^t(x), \delta^{(i)}(x') \right) = 0, \quad \forall i \]

Approach 3: Fully-Correlated Multi-Response

\[ y^t(x) = y^e(x) - \varepsilon = m^{(i)}(x) + \delta^{(i)}(x) \]

Assumption: Simulation models and the discrepancy functions follow the same spatial correlation function
Multi-model Fusion Procedure (Illustration of Approach 1)

**Covariance Calculation**

\[
\text{Cov} \left( y^{m(i)}(x), y^{m(j)}(x') \right) = e_i^T \Sigma^m e_j R^m(x, x')
\]

\[
\text{Cov} \left( y^e(x), y^{m(i)}(x') \right) = \rho^T \Sigma^m e_i R^m(x, x')
\]

**Data Integration via MVN (multivariate normal dist)**

\[
\left\{ \frac{1}{2} \right\} \exp \left( \frac{-1}{2T} \right) L(x) \sim_{MVN} \int \prod_{i=1}^{M} \left( \int \prod_{i=1}^{M} \right)
\]

**Hyperparameter Inference via MLE**

\[
L(\phi | d) \propto \left| V_d \right|^{-1/2} \exp \left\{ - (d - H\beta)^T V_d^{-1} (d - H\beta) / 2 \right\}
\]
The fidelity levels of simulator 1 and 2 are similar.

3 samples from Simulator 1, 7 samples from Simulator 2, 3 observations from experiment

<table>
<thead>
<tr>
<th>Approach</th>
<th>RMSE</th>
<th>u-pooling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approach 1</td>
<td>0.1530</td>
<td>0.0805</td>
</tr>
<tr>
<td>Approach 2</td>
<td>0.1636</td>
<td>0.0786</td>
</tr>
<tr>
<td>Approach 3</td>
<td>0.1573</td>
<td>0.0924</td>
</tr>
</tbody>
</table>
Validation Metrics

Root-mean-square error (RMSE)

\[ \text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{n}} \]

U-pooling

\[ u_i = F_{x_i}^m (y^e (x_i)) \]

---

The fidelity levels of both simulator 1 and 2 change over the design space.

- 5 samples from each simulator, 4 observations from experiment.
Example 2: Range-Dependent Model Fidelity

<table>
<thead>
<tr>
<th></th>
<th>Approach 1 with $0.01 &lt; \omega^2 &lt; 50$</th>
<th>Approach 2</th>
<th>Approach 3</th>
<th>Approach 1 with $0.01 &lt; \omega^2 &lt; 10$</th>
<th>Approach 1 with $0.01 &lt; \omega^2 &lt; 5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>0.5692</td>
<td>0.3329</td>
<td>0.3542</td>
<td>0.3598</td>
<td>0.2996</td>
</tr>
<tr>
<td>u-pooling</td>
<td>0.0797</td>
<td>0.0952</td>
<td>0.0966</td>
<td>0.1035</td>
<td>0.0823</td>
</tr>
</tbody>
</table>
Example 3: Fluidized-Bed Processes

- Used in the food industry to tune the effect of functional ingredients and additives.
- Important thermo-dynamic response: steady-state outlet air temperature.
- First studied by Dewettinck et al., 1999; employed by Reese et al., 2004; Qian et al., 2008.

- $V_f$: Fluid velocity of the fluidization air
- $T_a$: Temperature of the air from the pump
- $R_f$: Flow rate of the coating solution
- $P_a$: Pressure of atomization air
- $T_r$: Room temperature
- $H_r$: Room humidity

$Y^m_1$: Least accurate model because of its neglecting both heat losses and inlet airflow

$Y^m_2$: Intermediately accurate model taking those heat losses in the process

$Y^e$: Most accurate experiment test

Hierarchical Model Resources
Example 3: Fluidized-Bed Processes (Results)

<table>
<thead>
<tr>
<th></th>
<th>Approach 1</th>
<th>Approach 2</th>
<th>Approach 3</th>
<th>Qian and Wu’s approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>0.7402</td>
<td>0.6884</td>
<td>0.6925</td>
<td>/</td>
</tr>
<tr>
<td>u-pooling</td>
<td>0.1210</td>
<td>0.0706</td>
<td>0.1410</td>
<td>/</td>
</tr>
<tr>
<td>SRMSE</td>
<td>0.0177</td>
<td>0.0163</td>
<td>0.0169</td>
<td>0.020</td>
</tr>
</tbody>
</table>

A Multidisciplinary System

System Quantities of Interest (QOIs): $y_{sys}

Discipline 1

Discipline 2

Discipline $i$

System Analysis

$y_1$, Disciplinary outputs $y_2$

$u_{12}$, $u_{21}$

$y_i$, $u_i$

$Z$ is used to stand both disciplinary output $y_i$ and linking variables $u_i$. 

Aleatory Uncertainty  |  Epistemic Model Uncertainty
Multidisciplinary Statistical Sensitivity Analysis (MSSA)

VARIANCE-BASED SENSITIVITY INDICES

Impact of Aleatory Uncertainty

- \( \text{MSI}(X_i) = \frac{\text{Var}_X \left( \mathbb{E}_{Z,X_i} (Y|X_i) \right)}{\text{Var}(Y)} \)
- \( \text{TSI}(X_i) = 1 - \frac{\text{Var}_{Z,X_i} \left( \mathbb{E}_{X_i} (Y|Z,X_{-i}) \right)}{\text{Var}(Y)} \)

Impact of Epistemic Model Uncertainty

- \( \text{MSI}(Z_k) = \frac{\text{Var}_Z \left( \mathbb{E}_{Z,X} (Y|Z_k) \right)}{\text{Var}(Y)} \)
- \( \text{TSI}(Z_k) = 1 - \frac{\text{Var}_{Z_k,X} \left( \mathbb{E}_{Z_k} (Y|Z_{-k},X) \right)}{\text{Var}(Y)} \)

Challenges in SSA of model uncertainty

- Traditional Sobol’s method considers stochastic inputs as scalar variables
- \( Z \) are stochastic functional responses over model inputs.
- Nested situation where model uncertainty (\( Z \)) is a function of aleatory uncertainty (\( X \))
Separating Model Uncertainty in Disciplinary SRP

\[ y^e(x) = \hat{y}^e(x) + Z(x) \]

**DISCIPLINARY UNCERTAINTY QUANTIFICATION**

\[ u^e_i(x_i, x_s, u^e_i) = \hat{u}^e_i(x_i, x_s, u^e_i) + Z_{ui}(x_i, x_s, u^e_i) \]
\[ y^e_i(x_i, x_s, u^e_i) = \hat{y}^e_i(x_i, x_s, u^e_i) + Z_{yi}(x_i, x_s, u^e_i) \]

**TO AVOID NESTED SIMULATIONS IN SSA**

- Analytically derived multidisciplinary uncertainty propagation (MUA)
SRP-Based Multidisciplinary Uncertainty Analysis (MUA) Method

**System Analysis**

\[ u^c_i(x_i, x_s, u^c_i) = \hat{u}^c_i(x_i, x_s, u^c_i) + Z_{ui}(x_i, x_s, u^c_i) \]

\[ y^c_i(x_i, x_s, u^c_i) = \hat{y}^c_i(x_i, x_s, u^c_i) + Z_{yi}(x_i, x_s, u^c_i) \]

**Disciplinary Uncertainty Quantification**

- **Evaluation of means of linking variables and disciplinary outputs**
- **Evaluation of (co)variance of linking variables u**
- **Evaluation of (co)variance of disciplinary outputs y**
- **Evaluation of mean and (co)variance of system QOIs**

**A Matrix Form**

\[-A(u^c - \mu_u) \approx B(X - \mu_X) + Z_u, \quad y^c - \mu_y \approx (EA^{-1}B + F)(X - \mu_X) + EA^{-1}Z_u + Z_y\]

\[
\mu_{ui} \approx \hat{u}_i^c(\mu_{x_i}, \mu_{x_s}, \mu_{u^c_i}), \\
\mu_{yi} \approx \hat{y}_i^c(\mu_{x_i}, \mu_{x_s}, \mu_{u^c_i}), \\
\Sigma_{u} \approx (A^{-1}B)\Sigma_X(A^{-1}B)^T + (A^{-1})\Sigma_{u}(A^{-1})^T, \\
\Sigma_{y} \approx (EA^{-1}B + F)\Sigma_X(EA^{-1}B + F)^T + (EA^{-1})\Sigma_{Zu}(EA^{-1})^T + \Sigma_{Zy}.
\]
Case Study: An Aircraft Design Problem

System QOIs

$\$_{acq}$ - total acquisition cost

$\Delta t_{flight}$ - maximum time aloft

$A_{AV}$ - Ground area imaged by sensor

Design variables

Noise variables

Linking variables

Disciplinary outputs

System QOIs
Sensitivity Analysis for both Aleatory and Epistemic Uncertainties

Aleatory Uncertainty: 6~7%

Epistemic Uncertainty: 5%
OBJECTIVE

- To improve the *global* modeling capability of a multidisciplinary system

  Such that the epistemic uncertainty of system QOIs is acceptable over the input space.

**Resources**: Experiments and/or simulations

RESEARCH QUESTIONS

- **Where** in the input space of a multidisciplinary system shall we allocate more resources?

- **To what** disciplinary response(s) shall we allocate more resources?

- **Which** type of resource shall we allocate, experiments or simulations?
A Sequential Resource Allocation Strategy

1. Input Space Exploring
   - Generate samples over the input space
   - Assess the aggregated epistemic uncertainty of $Y_{sys}$ at each point
   - Evaluate whether uncertainty of $Y_{sys}$ is acceptable
     - Yes: END
     - No: Evaluate the impact of epistemic uncertainty from disciplinary responses at each point

2. Decision Making for Resource Allocation
   - Which (simulations vs. experiments)
   - What (disciplinary responses)
   - Where (sampling locations)

3. Updating Disciplinary Emulators

- SRP: Spatial-Random-Process
- MUA: Multidisciplinary Uncertainty Analysis
- MSSA: Multidisciplinary Statistical Sensitivity Analysis

$$\gamma(x_{ind}, x_s) \leq \frac{\sqrt{\text{Var}[Y_{sys}(x_{ind}, x_s)]}}{\int \int \|Y_{sys}(x_{ind}, x_s)\| dx_{ind} dx_s / \int \int dx_{ind} dx_s} \leq \alpha, \text{ for } \forall x_{ind}, x_s$$
AFTER SELECTING LOCATIONS AND RESPONSES...

Decision made in previous steps:
To allocate resources to selected $N_L$ locations for response $L$, and $N_{L'}$ locations for response $L'$, etc.

Suggest an affordable resource allocation plan
e.g., conducting experiments at $N_{Le}$ locations and simulations at $(N_L - N_{Le})$ locations for response $L$; similarly for $L'$, etc.

Monte Carlo loop
- Generate hypothetical data
- Update the emulators
- Evaluate the reduced uncertainty of $y_{sys}$

Evaluate the “expected” reduced uncertainty of $y_{sys}$

Is the reduced uncertainty of $y_{sys}$ acceptable?
- Yes: Use this plan
- No: Suggest another resource allocation plan
Case Study: Electronic Packaging


<table>
<thead>
<tr>
<th>$x_1$</th>
<th>Heat sink width ($m$)</th>
<th>$y_1$</th>
<th>Negative of watt density ($watts/m^3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_2$</td>
<td>Heat sink length ($m$)</td>
<td>$y_4$</td>
<td>Current in resistor #1 ($amps$)</td>
</tr>
<tr>
<td>$x_3$</td>
<td>Fin length ($m$)</td>
<td>$y_5$</td>
<td>Current in resistor #2 ($amps$)</td>
</tr>
<tr>
<td>$x_4$</td>
<td>Fin width ($m$)</td>
<td>$y_6$</td>
<td>Power dissipation in resistor #1 ($watts$)</td>
</tr>
<tr>
<td>$x_5$</td>
<td>Nominal resistance #1 at temperature 20 ºC ($Ω$)</td>
<td>$y_7$</td>
<td>Power dissipation in resistor #2 ($watts$)</td>
</tr>
<tr>
<td>$x_6$</td>
<td>Temperature coefficient of electrical resistance #1 ($°K^{-1}$)</td>
<td>$y_{11}$</td>
<td>Component temperature of resistor #1 ($°C$)</td>
</tr>
<tr>
<td>$x_7$</td>
<td>Nominal resistance #2 at temperature 20 ºC ($Ω$)</td>
<td>$y_{12}$</td>
<td>Component temperature of resistor #2 ($°C$)</td>
</tr>
<tr>
<td>$x_8$</td>
<td>Temperature coefficient of electrical resistance #2 ($°K^{-1}$)</td>
<td>$y_{13}$</td>
<td>Heat sink volume ($m^3$)</td>
</tr>
</tbody>
</table>
1st Iteration

- Model UQ: 40 experiments + 40 simulations
Selection of Input settings

1st Iteration
- Selection of input settings (from 2,000 samples) and responses

Resource Allocation

\[ \alpha = 10\% \]
Preposterior Analysis to Decide the Type of Resources to Allocate

DECISION (1\textsuperscript{ST} ITERATION)

(1) Allocate simulations of $y_{11}$ to points #2, 3, 10;
(2) Allocate simulations of $y_{12}$ to points #6~9;
(3) Allocate experiments of $y_{11}$ to points #1, 5;
(4) Allocate experiment of $y_{12}$ to point #4.
Subsequent Four Iterations (24 simulations + 10 experiments)
Summary

MODEL FUSION
- Approaches can handle both hierarchical and non-hierarchical rankings of fidelity
- Multiple approaches work equally well with reasonable assumptions

MULTIDISCIPLINARY UNCERTAINTY PROPAGATION AND SENSITIVITY ANALYSIS
- Considers both aleatory and epistemic uncertainties
- Utilizes the structure of SRP emulators, which allows for analytical derivation
- Decomposed disciplinary analyses, provide useful information for resource allocation

RESOURCE ALLOCATION FOR REDUCTION OF EPISTEMIC UNCERTAINTY
- Breaks a complex decision making problem into a sequential process
- Considers not only physical experiments but also simulations
Thank You!