A Bayesian Framework for the Integration of Separate and Integral Effects Validation Data

Nathan W. Porter  Postdoctoral Appointee

Vincent A. Mousseau  Principle Member of the Technical Staff
Motivation

• The complexity, quality, and quantity of data varies significantly for different types of experiments.

• Traditionally, experiments are organized into a hierarchy to account for this.

• Typical nomenclature: System, subsystem, component/benchmark, and unit.

• In traditional validation, these tiers are evaluated sequentially with no feedback.

• This avoids compensating errors.

• Here, we treat them consistently throughout the calibration and validation process.

Background and Theory
After code bugs and numerical errors have been minimized via SQA and verification, the remaining sources of uncertainty must be quantified.

- **Calibration** – parameter uncertainty and measurement errors
- **Validation** – model form errors
- **Prediction** – total uncertainty in the model

Here, we summarize a methodology which delineates these three processes in a consistent way while utilizing the calibration/validation pyramid.

1. Set up pyramid
2. Separate calibration and validation data
3. Iteratively perform calibration using Bayesian methods
4. Propagate parameter uncertainty through model for validation cases
5. Assess validation accuracy and/or predictive capability
• For calibration, statistical models generally assume that the experimental data is equal to some model with zero-mean Gaussian measurement noise.

\[ y_d = y_m(x, \theta) + \epsilon \]

• There are two primary choices for treatment of each parameter.
  1. A purely epistemic parameter has one “true” value which is unknown (e.g., physical constants). The posterior distribution represents epistemic uncertainty in the parameter value.
  2. A parameter may have combined aleatory and epistemic uncertainty. This is treated by assigning the parameter a probability distribution which represents the aleatory uncertainty, then the posterior density represents epistemic uncertainty in the distribution.

• Both methods can employ Bayesian Calibration to obtain estimates of the desired distributions, though they require different likelihood functions.\(^1\)

---

Calibration: Bayesian Methods

• In general, calibration is a statistical method to infer unknown parameter values/distributions by observing state variables and corresponding data (physical or computational experiments).

• Bayesian methods allow for the incorporation of prior information from previous experiments or expert knowledge.

• Solves Bayes’ formula, which formulates the desired posterior distribution in terms of the prior distribution and likelihood function.

\[
\pi(\theta | y) = \frac{\mathcal{L}(y|\theta)\pi_\theta(\theta)}{\int_{\Theta} \mathcal{L}(y|\theta)\pi_\theta(\theta)d\theta}
\]

• We employ sampling methods because (1) the denominator is difficult or impossible to integrate and (2) the product of the likelihood and prior cannot be easily sampled.
Calibration: Markov Chain Monte Carlo

- Methods which construct a sampling-based chain whose stationary distribution is equivalent to the desired posterior.
- Delayed Rejection Adaptive Metropolis (DRAM)
- Here, we use a hierarchical DRAM-within-Metropolis algorithm to statistically infer all unknown quantities.
- Priors are uninformative, with starting values assigned to an approximate maximum likelihood estimate (MLE).
Uncertainty Propagation

• In most cases, a Monte-Carlo propagation is used to approximate the effect of parameter uncertainty on model results.

• Requires a large number of samples to accurately predict distribution of the quantity of interest (QoI), and can therefore be computationally intensive.

• Alternatively, Wilks’ method\textsuperscript{1,2,3} can be used when comparison to safety or regulatory limits is required. This gives very little information about uncertainty or predictive capability.

---

Prediction: Validation Metrics

• Various methods to quantify the quality of code predictions in the presence of both measurement and prediction uncertainties.

• Frequentist and Bayesian hypothesis testing
  ◦ Calculate a p-value or Bayes factor, which indicate confidence in hypothesis.

• Reliability
  ◦ The probability that the observed difference is within a small interval $P(-\epsilon \leq D \leq \epsilon)$.
  ◦ Can be used to weight calibrated and “alternate” parameter distributions.

• Kolmogorov-Smirnov
  ◦ Maximum vertical distance between two CDFs, $0 \leq KS \leq 1$

• Area Metric
  ◦ Total area between two CDFs, positive

• Many others

Demonstration
A simple problem is selected to demonstrate the framework.

Friction and heat transfer in smooth tubes for turbulent flow.

The calibration pyramid has three components:
- Isothermal pressure drop experiments,
- Heat transfer experiments where pressure drop is not measured, and
- Simultaneous measurement of pressure drop and heat transfer.

Here, we perform only the calibration exercise, since propagation to validation cases is relatively self-explanatory.
Friction

• McAdams relation: \( f = 0.005 + 0.5Re^{-0.32} \)

• Experiments of pressure drop measurements in horizontal smooth tubes.

• Here, the statistical model is determined via initial frequentist analysis (minimization of AIC and BIC): \( \theta_1 \) is deterministic and \( \theta_2 \sim N(\mu, \sigma^2) \).

\[
f = \theta_1 + \theta_2 Re^{-0.32}
\]

<table>
<thead>
<tr>
<th>Year</th>
<th>Author</th>
<th>Pipe</th>
<th>Fluid</th>
</tr>
</thead>
<tbody>
<tr>
<td>1932</td>
<td>Nikuradse(^1)</td>
<td>Brass</td>
<td>Water</td>
</tr>
<tr>
<td>1982</td>
<td>Schlichting &amp; Gersten(^2)</td>
<td>Brass</td>
<td>Water</td>
</tr>
<tr>
<td>1998</td>
<td>Zagarola &amp; Smits(^3)</td>
<td>Aluminum</td>
<td>Air</td>
</tr>
<tr>
<td>2002</td>
<td>Swanson et al.(^4)</td>
<td>Stainless steel</td>
<td>Gases, He</td>
</tr>
<tr>
<td>2015</td>
<td>Furuichi et al.(^5)</td>
<td>Glass</td>
<td>Water</td>
</tr>
</tbody>
</table>

Heat Transfer

• Dittus-Boelter relation:

\[ Nu = 0.023 Re^{0.8} Pr^{0.4} \]

• Experiments of energy transfer to fluid flowing through heated smooth pipe.

• Statistical model: \( \theta_3 \) is deterministic and \( \theta_4 \sim N(\mu, \sigma^2) \).

\[ Nu = \theta_3 Re^{0.8} Pr^{\theta_4} \]

<table>
<thead>
<tr>
<th>Year</th>
<th>Author</th>
<th>Pipe</th>
<th>Fluid</th>
</tr>
</thead>
<tbody>
<tr>
<td>1928</td>
<td>Morris &amp; Whitman(^1)</td>
<td>Steel</td>
<td>Water &amp; oil</td>
</tr>
<tr>
<td>1931</td>
<td>Lawrence &amp; Sherwood(^2)</td>
<td>Copper</td>
<td>Water</td>
</tr>
<tr>
<td>1932</td>
<td>Sherwood &amp; Petrie(^3)</td>
<td>Copper</td>
<td>Water</td>
</tr>
</tbody>
</table>

Simultaneous Friction & Heat Transfer

• We use the data of Allen & Eckert for simultaneous calibration of the friction factor and Nusselt number.

• Only five data points; sparsity of the data is typical of many engineering problems.

• The plot shows the data along with McAdams and Dittus-Boelter relations.

• Likelihood is formulated as sum of friction and Nusselt number likelihoods.

• Again, statistical models are formulated via initial frequentist analysis with minimization of information criteria.

Simultaneous Friction & Heat Transfer

- First, we calibrate without using priors from separate effects data.
- Statistical model: Prandtl number is fixed, so Dittus-Boelter Prandtl exponent becomes unidentifiable. $\theta_1$ and $\theta_3$ are treated deterministically, $\theta_2 \sim N(\mu, \sigma)$.

$$f = \theta_1 + \theta_2 Re^{-0.32}$$
$$Nu = \theta_3 Re^{0.8} Pr^{0.4}$$

- Uncertainty in $\theta_4$ is not treated, Nusselt number uncertainty is dominated by zero-mean Gaussian error.
- Heat transfer result is calibrated only to these five data points, so the bias compared to Dittus-Boelter relation is not treated.
Simultaneous Friction & Heat Transfer

• Now, calibrate using separate effects results as priors.

• Statistical model: Dittus-Boelter Prandtl exponent becomes identifiable. $\theta_1$ and $\theta_3$ are treated deterministically, $(\theta_2, \theta_4) \sim N(\mu, \Sigma)$ with diagonal $\Sigma$.

\[
f = \theta_1 + \theta_2 Re^{-0.32} \\
Nu = \theta_3 Re^{0.8} Pr^{\theta_4}
\]

• Uncertainty in Nusselt is no longer dominated by zero-mean Gaussian error term.

• Heat transfer result correctly indicates that this data has a small positive bias.
Parameter Results

\[ f = \theta_1 + \theta_2 Re^{-0.32} \]

\[ Nu = \theta_3 Re^{0.8} Pr \theta_4 \]
Parameter Results

\[ f = \theta_1 + \theta_2 Re^{-0.32} \]

Parameters are significantly different when the separate effects priors are not included.
Parameter Results

\[ f = \theta_1 + \theta_2 Re^{-0.32} \]

\[ Nu = \theta_3 Re^{0.8} Pr \theta_4 \]

\( \theta_4 \) was unidentifiable before the addition of separate effects priors.
Parameter Results

Addition of separate effects priors reduces epistemic uncertainty, which is especially clear for deterministic parameters.

\[ f = \theta_1 + \theta_2 Re^{-0.32} \]

\[ Nu = \theta_3 Re^{0.8} Pr^{\theta_4} \]
Conclusion
Discussion

• Friction and heat transfer are empirical relations for the same physical process (boundary layers/turbulence), therefore the hyperparameters are not independent.
  ◦ Chilton-Colburn analogy
    \[
    \frac{Nu}{f} = 0.5 Re Pr^{1/3}
    \]
  ◦ In the future, this can be treated via a metropolis-within-DRAM hierarchical calibration, which allows accurate estimation of the hyperparameter covariance.¹

• Given data set(s) and corresponding state variables:
  1. Bayesian analysis is used to find parameter estimates using a hierarchy with priors.
  2. The parameter estimates are propagated through the validation problem(s).
  3. Quality of the predication is quantified.

Conclusion

• Calibration to separate effects data can remedy unidentifiability in the integral effects data due to model form, sparse data, or related physical processes.

• Separate effects data can also decrease epistemic uncertainty in models where integral data is sparse.

• It is important to not treat calibration/validation process or simulation model as “black boxes.”

• This process may break down or result in large uncertainties for cases with low quality or quantity of data.
  ◦ Large epistemic uncertainties indicate that data is sparse.
  ◦ Large aleatory uncertainty or noise indicate that the data is low quality.
  ◦ Can be used to direct future experimental work.

• Simultaneously calibrating to multiple datasets is mathematically equivalent to successive calibrations with priors. This process could also be incorporated into the framework.
Thanks to the following individuals for their assistance during the preparation of this work:

- Sandia National Labs: Joshua G. Mullins, Laura P. Swiler
- Los Alamos National Lab: Brian J. Williams
- North Carolina State University: Paul R. Miles, Ralph Smith
AIC and BIC

• Penalized-likelihood criteria that are often used to choose best predictor during regression analysis. Minimized value is the best fit.

• Measure of fit + penalty for complexity

• Akaike information criterion (AIC)

\[ AIC = -2 \ln \mathcal{L} + 2p \]

• Bayesian information criterion (BIC)

\[ BIC = -2 \ln \mathcal{L} + p \ln N \]

\( \mathcal{L} \): Likelihood

\( p \): number of parameters

\( N \): number of data points
Delayed Rejection Adaptive Metropolis (DRAM)

Design Parameters

- $n_s, \sigma_s^2, j_o, M$

Experimental Data

- $y_{i}, x_{i}$

$q_o = \arg \min_{\theta}[SS_{\theta}]$

Initial Optimization

- $s_0^2, SS_{\theta}, V$

Construct Candidate

$R_k = \text{chol} V$

$z_k \sim N(0, I_p)$

$\theta^* = \theta^{k-1} + R_k z_k$

$\alpha = \min \left(1, \frac{\pi(\theta^* | y)}{\pi(\theta^{k-1} | y)} \right)$

Accept Candidate

$\alpha > z_k$

$\pi(\theta^* | y)$

Second-stage candidate (DR)

$q^{*2} = q^{k-1} + \gamma_2 R_k z_k$

Update covariance (AM)

$V_k = s_p \text{cov}(\theta^0, ..., \theta^{k-1})$
Types of Distributions

- Maximum likelihood estimate (MLE): most accurate distribution, which excludes epistemic uncertainty.
- Unconditional distribution: includes both aleatory and epistemic uncertainty
  \[ f_\Theta(\theta) = \int_D f_\Theta(\theta|P = p)f_P(p) \, dp \]
- In this presentation, both aleatory and epistemic uncertainty are included in all distributions.
Friction Calibration

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>std</th>
<th>MC_err</th>
<th>tau</th>
<th>geweke</th>
</tr>
</thead>
<tbody>
<tr>
<td>\mu(\theta_4)</td>
<td>0.0059238</td>
<td>2.4415e-07</td>
<td>3.7964e-09</td>
<td>0.90063</td>
<td>1</td>
</tr>
<tr>
<td>\sigma(\theta_4)</td>
<td>0.48344</td>
<td>1.1076e-05</td>
<td>2.49e-06</td>
<td>706</td>
<td>0.9999</td>
</tr>
</tbody>
</table>

5/17/2019
Heat Transfer Calibration

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>std</th>
<th>MC_err</th>
<th>tau</th>
<th>geweke</th>
</tr>
</thead>
<tbody>
<tr>
<td>mu(\theta_2)</td>
<td>0.4971</td>
<td>0.0018767</td>
<td>4.5229e-05</td>
<td>1.1066</td>
<td>0.99963</td>
</tr>
<tr>
<td>sig(\theta_2)</td>
<td>0.030717</td>
<td>0.0013392</td>
<td>3.4161e-05</td>
<td>1.0807</td>
<td>0.99086</td>
</tr>
</tbody>
</table>

5/17/2019
Simultaneous Calibration

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>std</th>
<th>MC_err</th>
<th>tau</th>
<th>geweke</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu(\theta_2)$</td>
<td>0.482590</td>
<td>0.0009043</td>
<td>8.0967e-06</td>
<td>0.9714</td>
<td>0.99995</td>
</tr>
<tr>
<td>$\sigma(\theta_2)$</td>
<td>0.013545</td>
<td>0.0006455</td>
<td>6.865e-06</td>
<td>1.0015</td>
<td>0.99471</td>
</tr>
<tr>
<td>$\mu(\theta_4)$</td>
<td>0.497870</td>
<td>0.0018685</td>
<td>1.8028e-05</td>
<td>1.0028</td>
<td>0.99994</td>
</tr>
<tr>
<td>$\sigma(\theta_4)$</td>
<td>0.031149</td>
<td>0.0013237</td>
<td>1.4554e-05</td>
<td>1.0371</td>
<td>0.99896</td>
</tr>
</tbody>
</table>

- Figure 1: Scatter plots and kernel density plots for $\mu(\theta_2)$ and $\sigma(\theta_2)$.
- Figure 2: Scatter plots and kernel density plots for $\mu(\theta_4)$ and $\sigma(\theta_4)$.

5/17/2019