Statistical Validation of Scientific Models

"When I use a word," Humpty Dumpty said, in a rather scornful tone, "it means just what I choose it to mean - neither more nor less," *Through the Looking Glass*, Charles Lutwidge Dodgson (Lewis Carroll).
Examples of Scientific Accidents and their Reasons

Reference: Babuska, Nobile and Tempone, 2006

Modeling Error:

• *Tacoma Narrows Bridge*: Suspension bridge across the Puget Sound collapsed on November 7, 1940. Reason: Models did not adequately describe aerodynamic forces, the behavior of cables, or the effects of Von Karman vortices (http://www.youtube.com/watch?v=P0Fi1VcbpAI).

• *Hartford Civic Center Roof*: Collapsed on January 18, 1978. Reason: Joint models were inadequate.

• *Silo Failure*: Reason nonuniform forces due to soil configuration, grain flow, wind patterns. Necessitates granular flow model development.
Examples of Scientific Accidents and their Reasons

Numerical Error:

Examples of Scientific Accidents and their Reasons

Computer Science Error:

• *Explosion of Ariane 5*: June 4, 1996. Reason: Conversion from 64-bit floating point to 16-bit signed integer value exceed value that could be represented ([http://www.ima.umn.edu/~arnold/disasters/ariane.html](http://www.ima.umn.edu/~arnold/disasters/ariane.html)). Cost: $500 million.

[Image of Ariane 5 explosion]


Human Error:

Validation: The process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended model users.
Validation Metrics

- Experiment vs Model
  - 'Viewgraph' Norm

- Deterministic
  - Response vs Input

- Experimental Uncertainty
  - Response vs Input

- Numerical Error
  - Response vs Input

- Nondeterministic Computation
  - Response vs Input
Validation Metrics: Example

Metric 1:

Data

Model
Validation Metrics: Example

Metric 2:
Difficulties with Using Existing Experimental Data for Validation

1. Incomplete documentation of experimental conditions
   • Boundary and initial conditions
   • Material properties
   • Geometry of the system
2. Limited documentation of measurement errors
   • Lack of experimental uncertainty estimates

Note: It is critical that modelers, numerical analysts, and experimentalists work together

Reference: Oberkampf, Trucano and Hirsch 2004
Validation Strategies

Two Approaches:

1. Perform the experiment, independently, multiple times. If the experiments are truly independent, make estimates regarding the statistics of the uncertainty.
   -- Multiple, independent, validation experiments are often infeasible

2. Given estimates of pdf for parameters and input variables, determine how uncertainty propagates through full experimental process
   -- Termed `Validation Using Propagation of Uncertainty’
   -- Often less expensive
   -- Forces analysis of uncertainty in both models and experiments

Scientific Validation: Is the difference between model predictions and experimental observations significant relative to the uncertainty in the validation exercise?

Engineering Validation: Is the difference between model predictions and experimental observations significant relative to the uncertainty in the validation exercise, plan an acceptable error?
Validation Using Propagation of Uncertainty

Motivation: Consider the spring model

\[ m\ddot{u} + c\dot{u} + ku = F_0 \cos \omega t \]

Long-Term Behavior: Consider the oscillation amplitude

\[ U_0 = \frac{F_0}{\sqrt{(k - m\omega^2)^2 + (c\omega)^2}} \]

\[ \Rightarrow \frac{U_0}{F_0} = f(m, c, k, \omega) = \frac{1}{\sqrt{(k - m\omega^2)^2 + (c\omega)^2}} \]

Parameter and Measurement Densities: Assumed normally distributed

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( m )</td>
<td>1.00</td>
<td>0.001</td>
</tr>
<tr>
<td>( k )</td>
<td>2.00</td>
<td>0.05</td>
</tr>
<tr>
<td>( c )</td>
<td>0.25</td>
<td>0.05</td>
</tr>
<tr>
<td>( U_0/F_0 )</td>
<td>0.0</td>
<td>0.15</td>
</tr>
</tbody>
</table>
Validation Using Propagation of Uncertainty

Incorporation of Measurement Uncertainty:

- Measurements simulated by randomly choosing parameters and adding random measurement error
- Height of each error bar represents measurement uncertainty at 95% confidence level (±1.96σ)
- Does the model appear valid?
- Note: m = .9994, k = 1.9443, c = .1703
Validation Using Propagation of Uncertainty

Incorporation of Parameter and Measurement Uncertainty:

Recall: Let $X$ and $Y$ be random variables and consider the sum

$$Z = c_1 X + c_2 Y$$

Properties:

$$E(Z) = c_1 E(X) + c_2 E(Y)$$
$$\text{var}(Z) = c_1^2 \text{var}(X) + c_2^2 \text{var}(Y) + 2c_1c_2 \text{cov}(X, Y)$$

Present Problem: Assume experimental measurements are independent from estimates for model parameters

$$\sigma_{total}^2 = \sigma_{meas}^2 + \sigma_{pred}^2$$

Strategy: Use Monte Carlo simulations or sensitivity analysis to predict 95% confidence level for model predictions and add to standard deviation for measurement uncertainty.
Validation Using Propagation of Uncertainty

Incorporation of Parameter and Measurement Uncertainty:

Note: 19 of the 191 points do not lie within the 95% confidence band. Since this represents 9.95% of the points, does this mean that the model is not valid?
Validation Using Propagation of Uncertainty

Model Testing Using One Measurement:

- Take $U_0/F_0$ at $\omega = 1.45$ as a single test measurement
- Compute a histogram using 10,000 simulations
- Determine 95% confidence bounds

Conclusion: 5% chance of declaring model invalid when it is actually valid
Validation Using Propagation of Uncertainty

Model Testing Using Two Measurements:

- Take $U_0/F_0$ at $\omega = 1.0$ and $\omega = 1.45$ as a two test measurements
- Compute a histogram using 10,000 simulations
- Determine 95% confidence bounds

Hypothesis Test:

- $H_0$: Model is correct
- $H_1$: Model is incorrect
- $S_1$: $x \in 95\%$ confidence interval
- Test Statistic: Data values

Notes:

- Acceptance ranges of two measurements clearly interrelated. This is ignored in previous analysis.
- It is difficult to simulate, plot and interpret surfaces for large number of measurements.
- Alternative: Assume normally distributed
Validation Using Propagation of Uncertainty

Alternative Strategy:

• Assume that the PDF for the prediction uncertainty (including measurement uncertainty) is normally distributed.

• Use Monte Carlo simulations to estimate correlation between uncertainty in different predicted measurements.

• Recall: $\text{cov}(X, Y) = E[(X - E(X))(Y - E(Y))]$

• Estimate based on N samples:

$$\text{cov}(X, Y) = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \mu_x)(y_i - \mu_y)$$

Two Observations:

$$[\frac{U_0}{F_0}(1.0), \frac{U_0}{F_0}(1.45)] = [0.9701, 2.6545]$$

$$V = \begin{bmatrix} 0.0023 & 0.0008 \\ 0.0008 & 0.3546 \end{bmatrix}$$

Note: $1.96 \times \sqrt{0.3546} + 2.6545 = 3.8216$
Normally Distributed Prediction Uncertainty

Strategy: Contours of constant probability quantified by the following ellipses:

\[ r^2 = (x_1 - \mu_1, \ldots, x_n - \mu_n) V^{-1} \begin{bmatrix} x_1 - \mu_1 \\ \vdots \\ x_n - \mu_n \end{bmatrix} \]

Notes:

- Here each \( x_i \) represents a prediction of \( U_0/F_0 \) for the measurement frequency \( \omega_i \).

- Since prediction uncertainty is assumed normally distributed, \( r^2 \) is related to a \( \chi^2 \) distribution by the relation

\[ r^2 = \ell^2_{1-\alpha}(n) \]

where \( \ell \) is the value associated with the \( 100(1 - \alpha)\% \) confidence region for \( n \) parameters. These represent hypersurfaces of constant probability.
Normally Distributed Prediction Uncertainty

Recall: Let $X \sim N(\mu, V)$. Then

\[
f(x) = \frac{1}{\sqrt{(2\pi)^n |V|}} \exp \left[ -\frac{1}{2} (x - \mu)^T V^{-1} (x - \mu) \right]
\]

\[
= \frac{1}{\sqrt{(2\pi)^n |V|}} \exp \left[ -\frac{1}{2} r^2 \right]
\]

E.g., $V = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.01 \end{bmatrix} \Rightarrow r^2 = 10(x_1 - \mu_1)^2 + 100(x_2 - \mu_2)^2$

Let $C = V^{-1}$ and take

\[
C = \nu \lambda \nu^T
\]

where

\[
\lambda = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}, \quad \nu = \begin{bmatrix} v_{11} & \cdots & v_{1n} \\ \vdots & \ddots & \vdots \\ v_{n1} & \cdots & v_{nn} \end{bmatrix} = [v_1, \cdots, v_n]
\]

Note: $\nu^T = \nu^{-1}$
Normally Distributed Prediction Uncertainty

New Coordinate Vector:

\[ h = u^T (x - \mu) \text{ or } h_i = v_i^T (x - \mu) \]

\[ \Rightarrow r^2 = (x - \mu)^T \nu \lambda \nu^T (x - \mu) \]

\[ = h^T \lambda h \]

\[ = \sum_{i=1}^{n} \lambda_i h_i^2 \]

Second Transformation:

\[ z_i^2 = \lambda_i h_i^2 \]

\[ \Rightarrow r^2 = z^T z = z_1^2 + \cdots + z_n^2 \]

Probability that Point Lies Inside Hypersphere:

\[ P(r^2 \leq \ell^2) = \int \cdots \int \frac{1}{\sqrt{(2\pi)^n}} e^{-r^2/2} dz_1 \cdots dz_n \]

Note: \[ |V|^{-1/2} = |C|^{1/2} = (\lambda_1 \cdots \lambda_n)^{1/2} \]
Normally Distributed Prediction Uncertainty

Note: The volume element can be expressed as

\[ dV = \frac{n\pi^{n/2}r^{n-1}dr}{\Gamma(n/2 + 1)} \]

Thus

\[ P(r^2 \leq \ell^2) = \frac{n2^{-n/2}}{\Gamma(n/2 + 1)} \int_0^\ell e^{-r^2/2}r^{n-1}dr \]

Note: Transformation \( r^2 = x \) yields integral of \( \chi^2 \) density with \( n \) DOF

\[ \text{e.g., } n = 1: \quad P(r^2 \leq \ell^2) = \sqrt{\frac{2}{\pi}} \int_0^\ell e^{-r^2/2}dr = \text{erf}(\ell^{2-\ell/2}) \]

\[ n = 2: \quad P(r^2 \leq \ell^2) = \int_0^\ell e^{-r^2/2}rdr = 1 - e^{-\ell^2/2} \]

Summary:

\[ r^2 = (x - \mu)^TV^{-1}(x - \mu) = \ell^2_{1-\alpha}(n) \]
Normally Distributed Prediction Uncertainty

Example: 2 Measurement Points

$$\ell^2_{0.95}(2) = (2.447)^2 = 5.991$$

$r^2$ value for data point with 10,000 simulations: 3.7781
Normally Distributed Prediction Uncertainty

**Example:** 191 Measurement Points

\[ \ell^2_{0.95}(191) = 224.2 \]

\( r^2 \) value for data points with 10,000 simulations: 11.89

**Numerical Issues:** The covariance matrix \( V \) can be very ill-conditioned. This necessitates use of the Moore-Penrose or pseudoinverse \( \text{pinv}.m \)
Alternative Method: Optimization-Based

Notes:

• Avoids the assumption that model uncertainty is normally distributed
• Requires statistical model for input parameters and measurement uncertainty
• Example: Consider first $\omega = 1.45$ and uncertainty in $c$
Optimization Method

**Assumption:** Measurement uncertainty independent from parameter uncertainty which yields

\[
    r^2 = [c - \bar{c}, \frac{U_0}{F_0(1.45)} - \text{meas}(1.45)] V^{-1} \begin{bmatrix} c - \bar{c} \\ \frac{U_0}{F_0(1.45)} - \text{meas}(1.45) \end{bmatrix}
\]

**Test:** Reject model as invalid at 95% confidence level if there is no value of c that yields model prediction passing through confidence region

\[
    \left( \frac{c - \bar{c}}{\sigma_c} \right)^2 + \left( \frac{U_0/F_0 - \text{meas}(U_0/F_0)}{c_p} \right)^2 = \ell^2_{0.95}(2) = r^2_{\text{crit}}
\]

**Distance Function:**

\[
    \left( \frac{c - \bar{c}}{\sigma_c} \right)^2 + \left( \frac{U_0/F_0 - \text{meas}(U_0/F_0)}{c_p} \right)^2 = r^2
\]

**Notes:**

- **Parameters:** c and meas(U_0/F_0) at \( \omega = 1.45 \)
- **Here** \( \ell^2_{0.95} = (2.447)^2 \)
Optimization Method

**Strategy:** Minimize $r$ and see if resulting value is less than 2.447.

**Conclusion:** We cannot declare model invalid at 95% confidence level since there are values of $c$ that yields values of $r < 2.447$. 
Optimization Method

Multiple Measurements and Parameters: e.g., consider 2 measurements and 3 parameters

Minimize

\[ r^2 = d^T V^{-1} d \]

where

\[ d^T = [m - \bar{m}, \ k - \bar{k}, \ c - \bar{c}, \ U_0/F_0(1.45) - \text{meas}(1.45), \ U_0/F_0(1.0) - \text{meas}(1.0)] \]

\[
V = \begin{bmatrix}
0.001^2 & 0 & 0 & 0 & 0 & 0 \\
0 & 0.05^2 & 0 & 0 & 0 & 0 \\
0 & 0 & 0.05^2 & 0 & 0 & 0 \\
0 & 0 & 0 & 0.015^2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0.015^2 & 0
\end{bmatrix}
\]

Note: \( \chi^2_{0.95}(5) = 11.07 \Rightarrow r_{\text{crit}} = 3.33 \)

Note: Optimal values \( m = 1.000, \ k = 1.9456, \ c = 0.1703 \) yields \( r = 1.9313 \)
Optimization Method

Multiple Measurements and Parameters: e.g., consider 191 measurements and 3 parameters

Note: $\chi^2_{0.95}(194) = 227.5 \Rightarrow r_{crit} = 15.08$

Note: Values $m = 1.000, k = 1.9456, c = 0.1703$ yields $r = 2.0257$