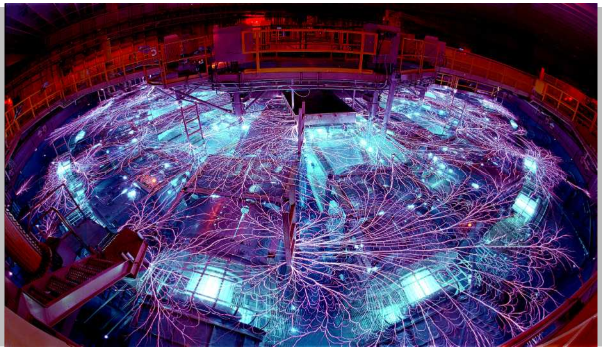
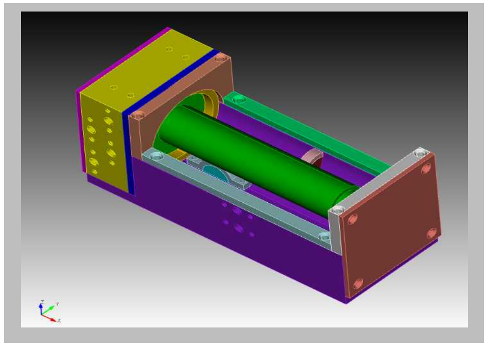
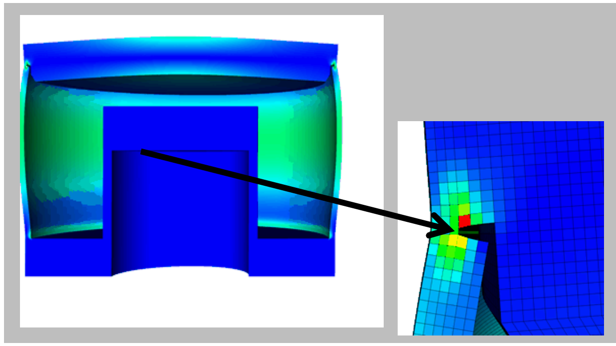
SAND2013-2760P

Exceptional service in the national interest

Sandia National Laboratories







Validation Hierarchy and Aggregation to Target Application (Roll-up)

Richard Hills

SAND Number:

U.S. DEPARTMENT OF ENERGY

NNSA National Nuclear Security Administration

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Outline

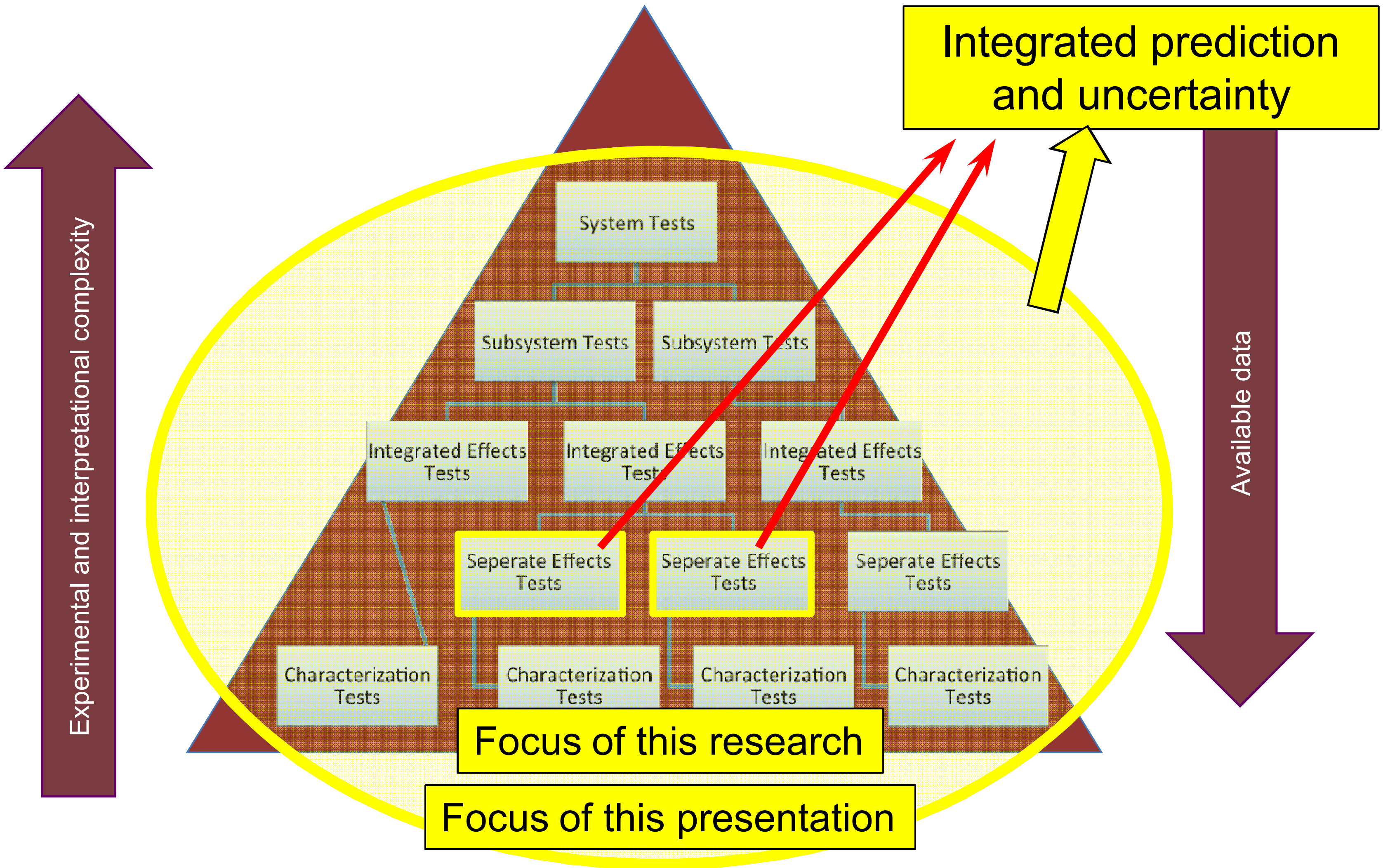
- Background
  - Validation hierarchy and aggregation to an application
  - Other approaches
- Present Approach
  - The concept
  - Distance between experiment and application
  - Aggregation/roll-up across hierarchy
  - Other uses
- Examples

High priority uncertainties for aggregation to a target application

- Model parameter uncertainty
- Mesh convergence and stochastic convergence uncertainty
- Model form uncertainty
  - Alternative models
  - Evidenced by validation experiments
    - Observed differences and their uncertainties
    - Incomplete validation coverage of application

Focus of this work

Validation Hierarchy or Pyramid



Why do we care?

- We sometimes can't run experiments at the conditions of the application, but can run a set of experiments at more benign conditions.
  - What is the impact on target application predictions?

- We can observe model form error from experiments lower in the hierarchy
  - Are these differences really important to the application?
  - If so, what can we do?

To address these questions, we need to develop a relationship between the validation experiments and the application

Other Approaches to Roll-Up

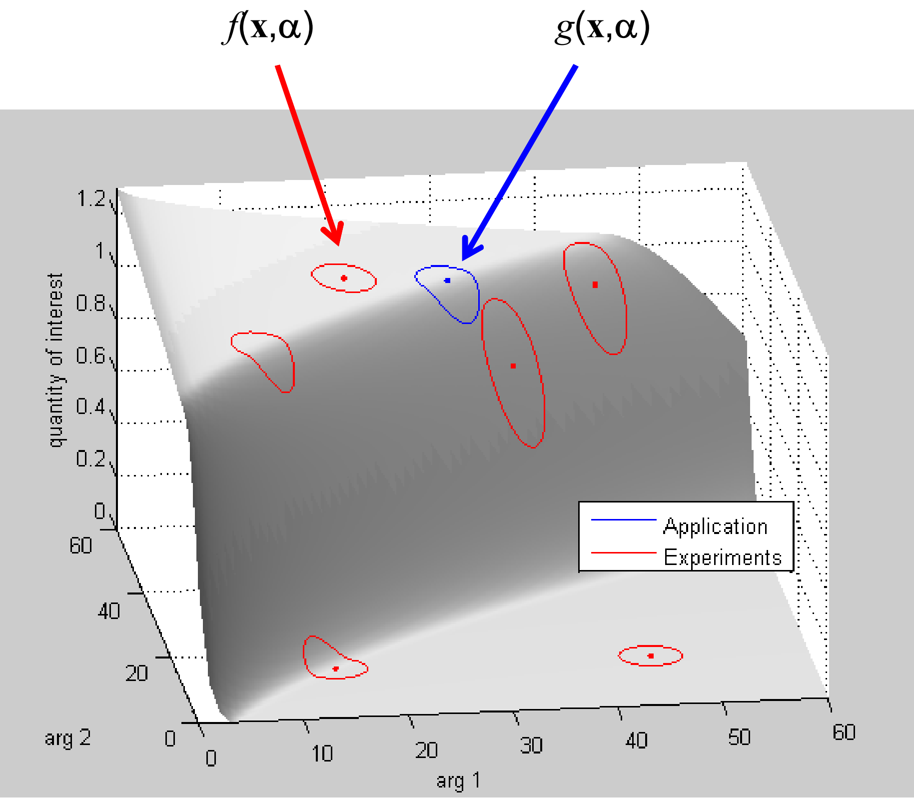
- Calibration or multiple calibrations (Babuška, et al.)
  - Sensitivity to calibrated predictions to multiple calibration data set
  - Assumes that the effect of model form error can be captured by multiple calibrations
- Calibration including model deficit term (Kennedy and O'Hagan)
  - Useful when validation measurement types are the same as the response quantities of interest for the target application (homogeneous hierarchy)
  - Gaussian Process Models are often used for model deficit term – does not preserve original conservation principles
- Bayesian net – evaluates a measure of reliability based on validation results and propagates to target application through common parameters (Mahadevan)
  - Experimental results from different location in the validation hierarchy weighted equally
  - No model deficit term in present configuration

Present Approach

- Present approach – develops a meta-model to relate validation experiments to application for a heterogeneous hierarchy
  - Accounts for 'distances' between validation experiments and target application
  - Quantifies completeness of validation hierarchy
  - Can be used to develop a model deficit term that allows for mixed variable types
- Best approach? – open research question
- Issue:
  - All approaches utilize the CompSim models for the validation experiments and the target application – if physics is missing, all approaches are approximate at best, misleading at worst

Our approach to the hierarchy

- Define a ‘distance’ measure in terms of <u>relative behavior </u>of the <u>models</u> of experiments to the <u>model</u> of application
  - Use a distance based on behavior, not on differences in arguments
  - Use non-linear sensitivity analysis relative to the model arguments to characterize relative behavior
- Aggregate or weight the suite of experimental <u>models</u> to ‘best represent’ the target application <u>model</u> by minimizing the distance between the aggregated model and the target application model such that the
  - Aggregated model has the same sensitivity to important arguments as the application model



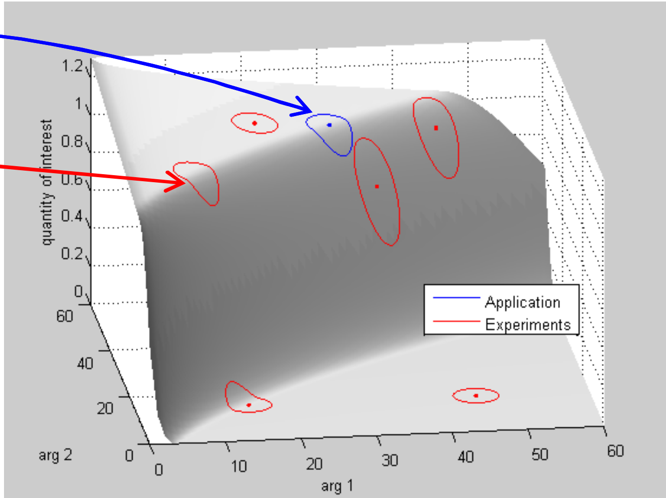
Arguments – $(\mathbf{x},\alpha)$

Key assumption 1

Distance and Aggregation

Characterize Behavior: Latin Hypercube Sampling over neighborhoods

Application: $\Delta\mathbf{g} = \begin{bmatrix} \Delta g_1 \\ \vdots \\ \Delta g_m \end{bmatrix}$; Validation Exp: $\Delta\mathbf{f} = \begin{bmatrix} \Delta f_1 \\ \vdots \\ \Delta f_m \end{bmatrix}$



Aggregation:

$$\Delta\mathbf{g} = \begin{bmatrix} \Delta g_1 \\ \vdots \\ \Delta g_m \end{bmatrix};\ \Delta\mathbf{F} = \begin{bmatrix} \Delta f_{11} & \cdots & \Delta f_{1n} \\ \vdots & \ddots & \vdots \\ \Delta f_{m1} & \cdots & \Delta f_{mn} \end{bmatrix};\ \mathbf{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$$

Key assumption 2 → $\Delta\mathbf{g} \cong \Delta\mathbf{F}\,\mathbf{w}$

Key difficulty: $\Delta\mathbf{F}$ often algorithmically singular

Columns of $\Delta\mathbf{F}$:

- Can represent different experiments, quantities, or times from same experiment
- Consistent sampling across neighborhoods required

Modified Partial Least Squares Regression used to find w

Partial Least Squares Regression (PLSR)

- Other uses of PLSR
  - Originally used in the social sciences
  - Heavily used in chemometrics
  - Also use in anthropology, neuroscience, sensometrics, bioinformatics
- PLSR finds the multidimensional directions in the measurement space that explains the maximum multidimensional variance directions in the prediction space.
- Well suited when more weights w (or measurements) than the number of LHS samples and multi-colinearity amongst the $\Delta\mathbf{f}$ vectors

Latent Variables

- Develops an intermediate space in terms of p latent variables and the associated directions where $p \leq \text{rank}(\Delta\text{F})$
- For present application, p is an effective rank and is related to
  - The number of model arguments that have a significant effect on the predictions (increases p)
  - The amount of uncertainty in the validation and application model parameters (increases p)
  - The amount of measurement uncertainty (decreases p)
- Only need a sufficient number of LHS samples to estimate the p latent variables

Basic Assumption

Predictions from models with similar behavior will have similar model form error to first order

- Perhaps more appropriate when model form error is a secondary effect
- Need judgment as to the validity of the basic assumptions for a specific application

There is no magic bullet and one should proceed with caution!

Products and Process

Sandia National Laboratories

- Importance ordering of various measurements to application

$$r = \sqrt{(\Delta\mathbf{g} - \Delta\mathbf{f}\,\mathrm{w})^{\mathrm{T}}(\Delta\mathbf{g} - \Delta\mathbf{f}\,\mathrm{w})}$$

decreasing r implies increasing importance

- Completeness: Is validation hierarchy sufficiently complete?

$$\text{Coverage Residuals: } \mathbf{r} = \Delta\mathbf{g} - \Delta\mathbf{F}\,\mathbf{w}$$

Step 1

- Validation metric: are validation differences significantly small for to application?

$$r^2 = \mathbf{w}^{\mathrm{T}}\,(\boldsymbol{\gamma} - \mathbf{f})^{\mathrm{T}}\,\mathbf{cov}(\boldsymbol{\gamma} - \mathbf{f})^{-1}\,(\boldsymbol{\gamma} - \mathbf{f})\,\mathbf{w}$$

Step 2

- Project differences to application with caution

$$\Delta\mathbf{g}_{\mathbf{proj\ meas}} \cong \Delta\mathbf{F}_{\mathrm{meas}}\,\mathbf{w} = (\boldsymbol{\gamma} - \mathbf{f})\,\mathbf{w} = \begin{bmatrix} \gamma_1 - f_1 \\ \vdots \\ \gamma_n - f_n \end{bmatrix}\mathbf{w}$$

Step 3

Process

Simple Examples

$$\frac{\partial u}{\partial t} + c\frac{\partial u^p}{\partial x} = d\frac{\partial^2 u}{\partial x^2}$$

c = 0, $d \neq$ 0: diffusion
$c \neq$ 0, d = 0, p = 1: convection
$c \neq$ 0, d = 0, p = 2: Burgers
c, $d \neq$ 0, p = 1: conv-diff
c, $d \neq$ 0; p = 2: diff. Burgers

$$u(x,0) = \begin{cases} 0.5; & x < 1 \\ x - 0.5; & 1 \leq x < 2 \\ 1.5; & 2 \leq x < 3 \\ 4.5 - x; & 3 \leq x < 4 \\ 0.5; & 4 \leq x \end{cases}$$

$$u(0) = u(20)$$

Uncertainties considered:

- model parameter for validation experiments and application
- measurement

Example 1: Completeness

- We can’t afford to do validation experiments for all of the physics!

- What is the impact of not doing validation experiments related to the non-linear feature of Burgers’ equation?

Coverage

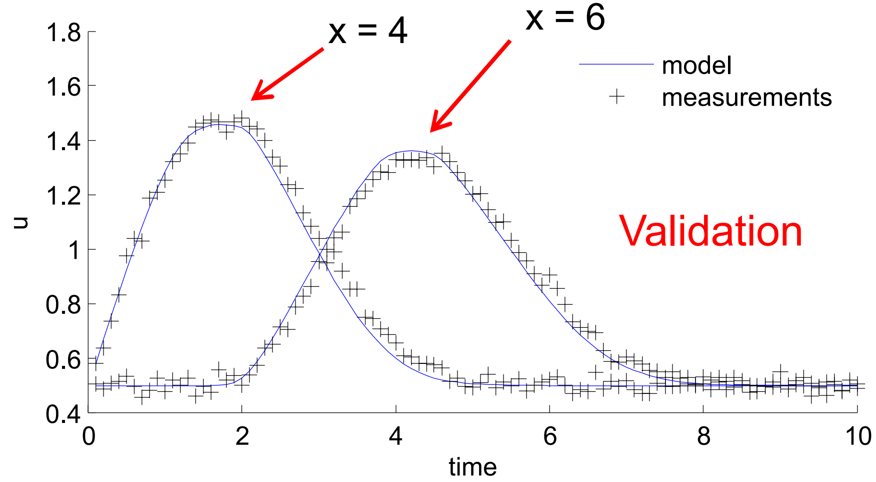
Application Physics:
- Diffusive Burgers'

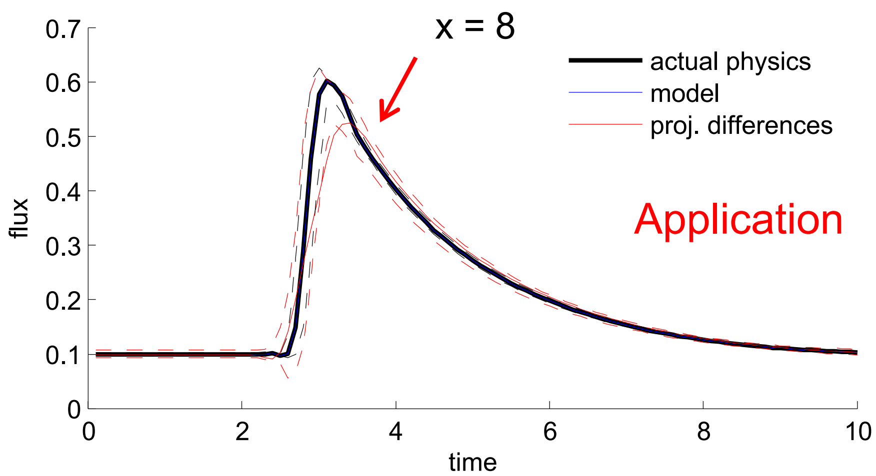
Validation Exp. Physics
- Convection-Diffusion

Measured
- u at x = 4, 6
- Independent experiments

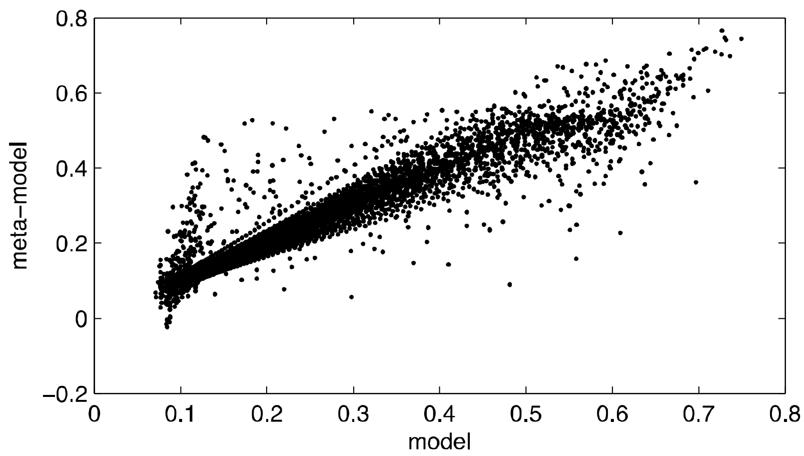
Application quantity of interest
- flux at x = 8

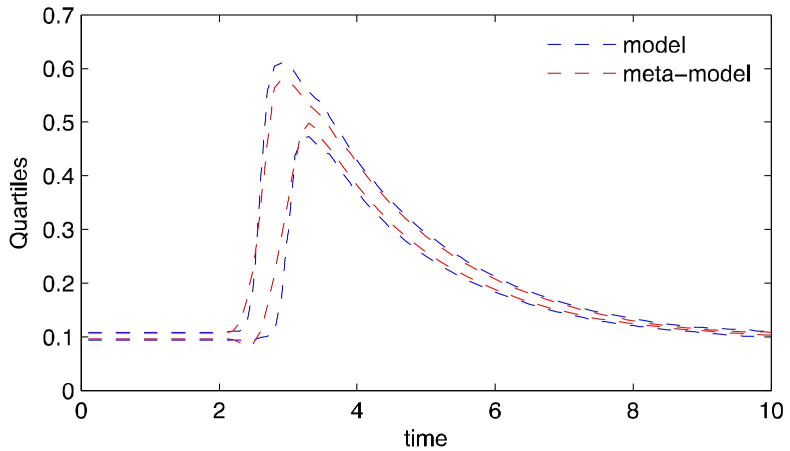
Notes
- Uncertainty in model parameters for both validation experiments and application (specified)
- Measurement uncertainty (specified)





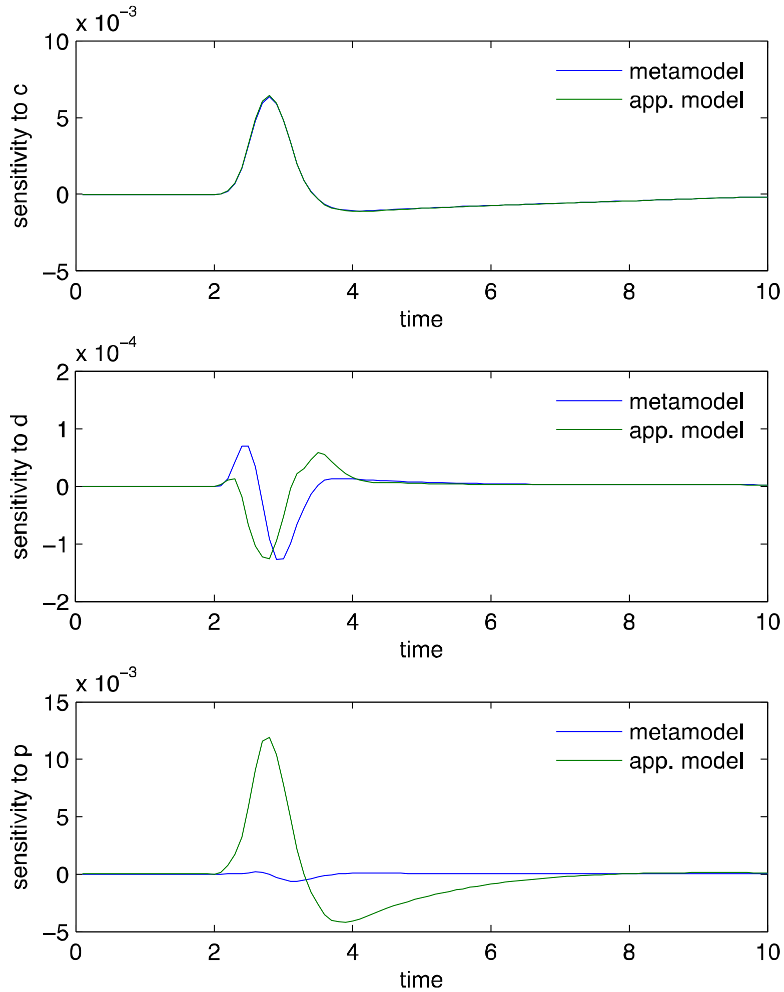
Coverage: Assessment





Application not fully resolved by experiments

Coverage: Sensitivity



Sensitivity to linear convective term: Well represented

Sensitivity to diffusive term: Phase change, but magnitudes well represented

Sensitivity to non-linear Burgers' term: Dominant sensitivity, not represented

Require experiments to test non-linear term

Example 2: Observed Model Form Error

- Our validation experiments show strong model form error!
- Calibration results in parameters well outside the expected range!
- Can we use the observed validation differences in u to correct the predictions of flux for the application?
- Will some measurements of flux help in the correction at the application level?

Very hard test problem! As example 1 showed, the missing physics are the dominant physics!

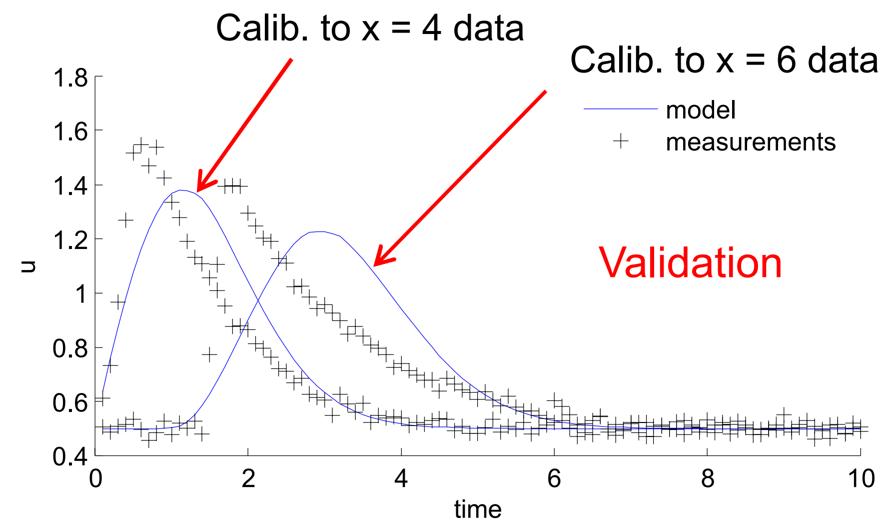
Observed Model Form Error in u: Predicting flux

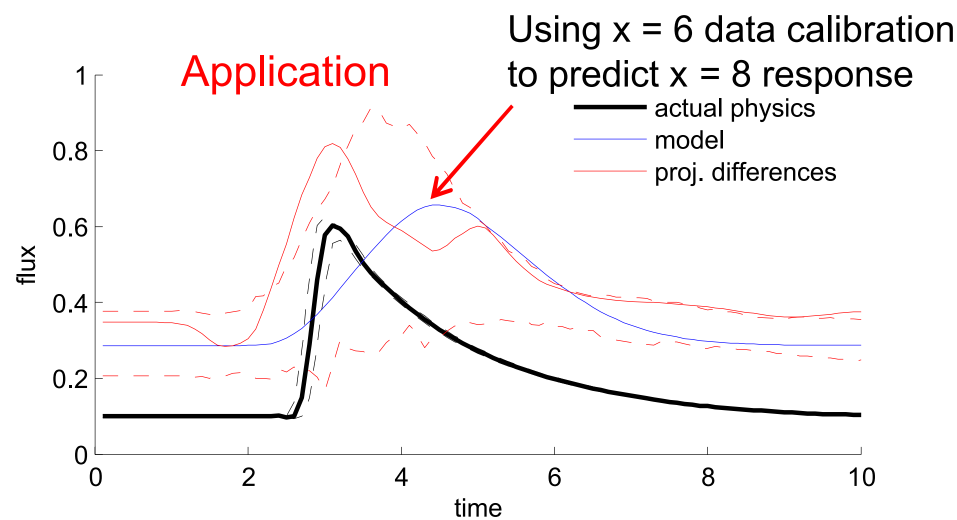
Actual Physics:
- Diffusive Burgers'

Modeled Physics
- Convection-Diffusion

Measured
- u at x = 4, 6
- Independent experiments

Application quantity of interest
- flux at x = 8





Correcting flux predictions using observed model form error in u performs poorly: other experimental evidence needed!

Some Flux Measurements

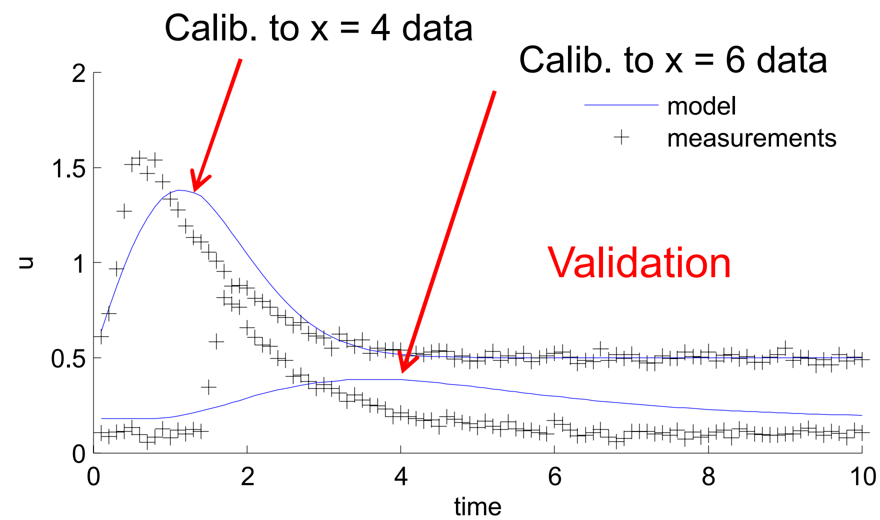
Sandia National Laboratories

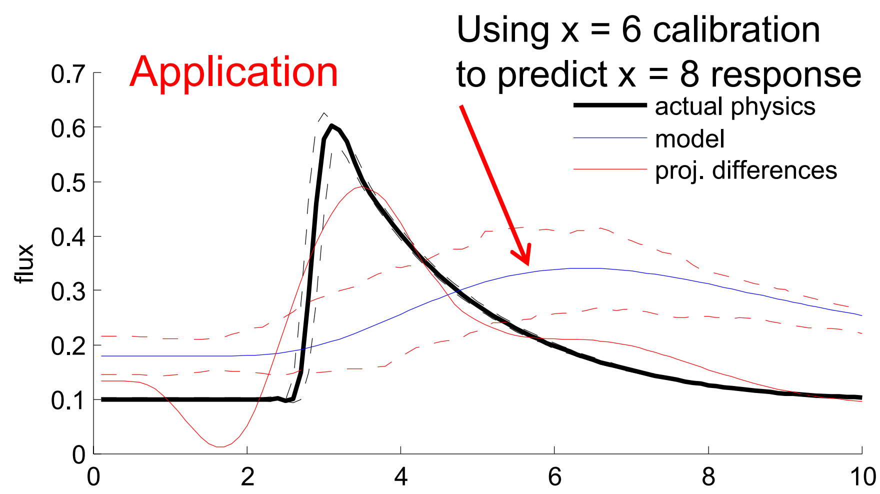
Actual Physics:
- Diffusive Burgers'

Modeled Physics
- Convection-Diffusion

Measured
- u at x = 4, flux at x = 6
- Independent experiments

Application quantity of interest
- flux at x = 8





Flux measurements help for the projected differences, but not for calibration

Projected differences well outside quartiles indicates that significant physics is missing from application prediction: should further develop model

Recap: What are we really doing?

- Using the models, not the observed validation differences, to develop a relationship between the experimental measurements and the application predictions
- Applicable to fully heterogeneous validation hierarchies
- Provides insight into the relationship between the validation measurements and the application for experimental design
- Can use this relationship to map the observed validation differences to the application

Discussion

- Issue: Present approach can project observed model form due to residuals from a calibration, but can also display non-physical behavior (i.e. oscillations) when model form error is large.
  - Are there better ways to evaluate the weights given the singular system?
    - PLS structural modeling, non-linear and constrained PLS
- Are there features of this approach that can be adapted to other approaches?
  - For example, can the concept of distance between experiment and application be used with the Bayesian net approach discussed earlier?
- Much research to be done in the development and choice of the best methodology!
- To be useful to SNL, the methodology must be able to address heterogeneous validation hierarchies

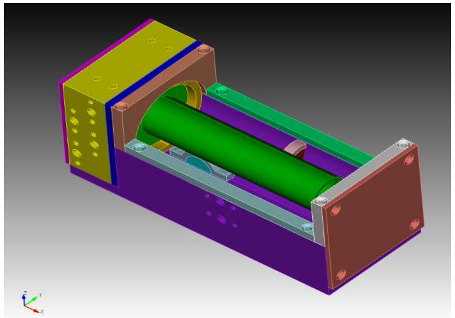
Methodology cannot create knowledge out of nothing!

Judgment is required to cast a sufficiently broad net in the validation hierarchy and to decide when additional model building is appropriate.

Next Steps

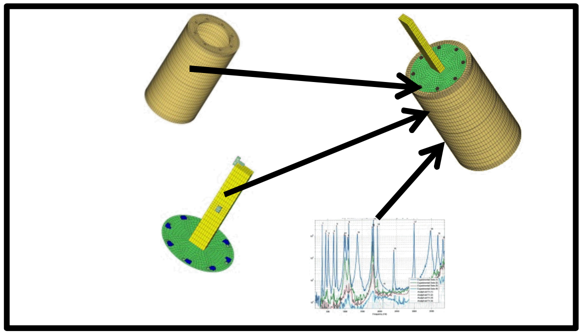
In progress:

- Test methodology for 2 SNL applications
  - System-Generated Electro-Magnetic Pulse
  - Re-entry structural response



Future plans

  - FY14: Additional development of methodology
  - FY16 Thermo-mechanical breech L1 if appropriate



Questions?

Relevant Publications

Other Basic Approaches:

- Babuska, I., F. Nobile, R. Tempone (2008), A Systematic Approach to Model Validation Based on Bayesian Updates and Predicted Related Rejection Criteria, Computer Methods in Applied Mechanics and Engineering 197:2517-2539.
- Kennedy, M.C., O'Hagan, A. (2001), Bayesian Calibration of Computer Models, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 63(3):425-464.
- Mahadevan, S. (2011), Roll-Up of Multi-Level UQ Activities towards System Level QMU, Presented at Sandia National Laboratories, Albuquerque, NM, Oct. 4.

Current Approach

- Hills, R. G., I. H. Leslie (2003), Statistical Validation of Engineering and Scientific Models: Validation Experiments to Application, SAND2003-0706, Sandia National Laboratories, Albuquerque, NM.
- Hills, R.G. (2006), Model Validation: Model Parameter and Measurement Uncertainty, Journal of Heat Transfer 128, pp. 339-351.
- Hills, R.G., J.R. Hamilton (2009), Validation Experiments to Application: A Model Based Approach, SAND2009-1091 Unlimited Release Printed February 2009
- Hamilton, J. R., R. G. Hills (2010a), Relation of Validation Experiments to Applications, Numerical Heat Transfer, Part B: Fundamentals, 57: 5, pp. 307-332.
- Hamilton, J. R., R.G. Hills (2010b), Relation of Validation Experiments to Applications: A Nonlinear Approach, Numerical Heat Transfer, Part B: Fundamentals, 57: 6, pp. 373-395.
- Hills, R. G. (2013), Roll-up of Validation Results to a Target Application, Sandia Report, in revision