VERIFICATION AND VALIDATION:

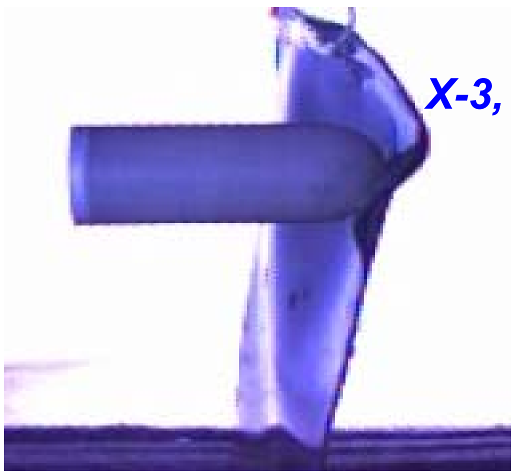
QUANTIFYING PREDICTION UNCERTAINTY AND DEMONSTRATING SIMULATION CREDIBILITY

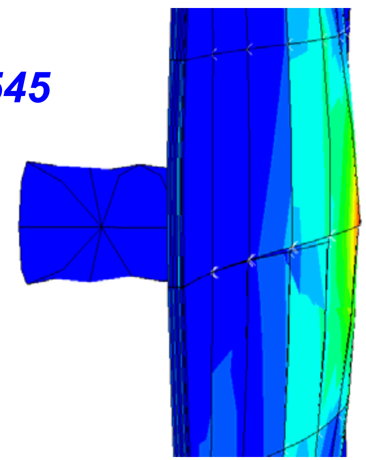
François M. Hemez

Member of the Technical Staff in X-Division

Los Alamos National Laboratory
X-3, Mail Stop F644, Los Alamos, New Mexico 87545

hemez@lanl.gov





This presentation has been approved for unlimited, public release.
Date: May 5, 2008
Reference: LA-UR-08-2849
Level: Unclassified

Abstract

VERIFICATION AND VALIDATION: QUANTIFYING PREDICTION UNCERTAINY AND DEMONSTRATING SIMULATION CREDIBILITY

Verification and Validation (V&V) refers to a broad range of activities that are carried out to provide evidence that measurements and predictions are credible and scientifically defendable. This presentation offers an introduction to the main concepts of V&V and lessons learned after fifteen years of research, development, and application of V&V technology at the Los Alamos National Laboratory (LANL). The discussion is somewhat restricted to Structural Dynamics even though V&V at LANL reaches across software quality assurance, verification, data analysis and archiving, engineering simulation, computational physics and astrophysics simulation, and the quantification of uncertainty. While high-level concepts are emphasized, references are made available for the implementation of specific tools or application case studies. The cornerstone of V&V is threefold with, first, showing whenever possible that predictions of numerical simulations are accurate relative to test data over a range of settings or operating conditions; second, quantifying the sources and levels of prediction uncertainty; and, third, demonstrating that predictions are robust, that is, insensitive, to the modeling assumptions and lack-of-knowledge.

(Approved for unlimited, public release on May 5, 2008, LA-UR-08-2849, Unclassified.)

Disclaimer

- The opinions expressed in this material are mine and do not necessarily reflect those of line-management at the Los Alamos National Laboratory (LANL) or U.S. Department of Energy.
- This presentation is a high-level introduction to the main concepts of Verification and Validation (V&V); it purposely-so contains little-to-no technical detail.
- Much of the material is extracted from a graduate-level course, that I first taught at the University of California San Diego (UCSD) in 2005; and a short-course version for industry training, that has been taught since 2001.

Who Am I?

- **I have been a Technical Staff Member at Los Alamos since 1997; currently with X-Division where I manage the Code Verification project.**
- **I have worked on many Verification and Validation (V&V) studies applied to structural health monitoring, material modeling, weapon physics and engineering.**

François Hemez, LANL

- **Currently lead the Code Verification effort of physics performance codes developed at Los Alamos.**
- **Teach V&V at UCSD since 2005.**
- **Have authored 21 peer-reviewed manuscripts and 270+ other papers and reports since 1992; most of which related to V&V one way or another.**

Main Take-home Points

- Verification and Validation (V&V) is a rigorous and scientifically defendable procedure by which analysts can demonstrate the *credibility* of their predictions.

- The overarching goal is to identify *all significant* sources of uncertainty and lack-of-knowledge, and *quantify* the effects that they have on predictions.

$$\text{“Confidence} \approx \frac{1}{\text{Total Uncertainty}}\text{”}$$

- Credibility is demonstrated, not just by showing that numerical predictions match physical measurements, but understanding the extent to which predictions are *robust* (or insensitive) to modeling assumptions.

Outline

- Definitions and language of V&V
- An example in Structural Dynamics
- High-level description of a few V&V activities
- Two lessons (often learned the hard way)
- Overview of how V&V is structured at Los Alamos
- Closing remarks

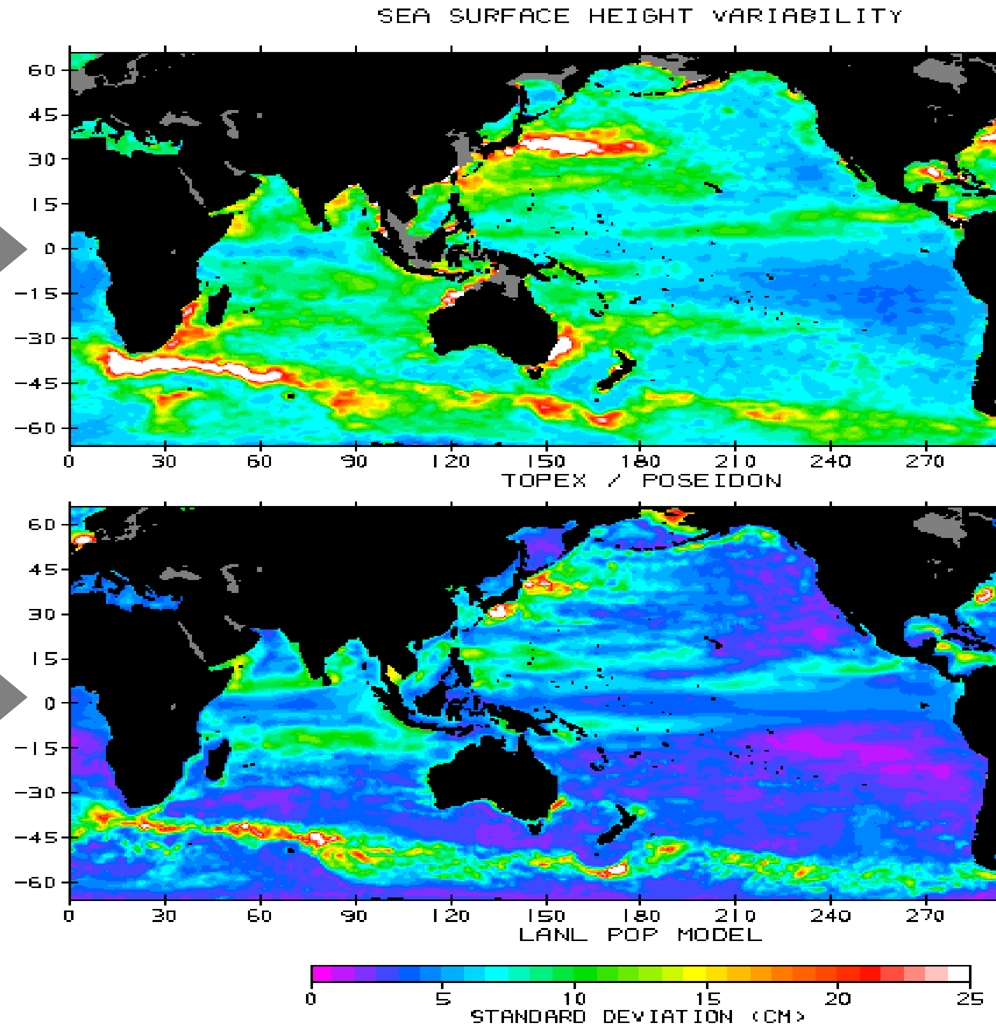
When is V&V Relevant?

- **Are you managing a code project?**
- **Are you in an environment where it is essential to develop, or invest in, a predictive capability?**
- **Do you operate in a product maintenance mode where questions asked by customers are often answered through numerical simulations?**
- **Do you have to support high-consequence decisions by examining a combination of experimental evidence, current knowledge, and numerical simulations?**
- **If the answer to one of these questions is *"yes,"* then V&V at some level or another, makes sense because it is how you establish *credibility*.**

Let's Start … Is This *"Validation"*?

Observation of the TOPEX/Poseidon Satellite

Calculation of the LANL/POP Simulator (~ 2002)



The POP/TOPEX Simulation

- The simulation couples models of atmosphere, ocean, and ice cap with a ½-degree grid resolution and ≈ 20 non-uniform zones through the atmosphere/ocean.(#)
- Few laboratory experiments (if any?) are available to understand the coupling between atmosphere, ocean, and ice. As a result the calculation is weakly coupled.
- Are the conservation laws verified? Has the advection based mapping of various fields been verified?
- Is a resolution of ≈ 20 grid points through the depth of the atmosphere/ocean enough to calculate discrete solutions in the asymptotic regime of convergence?

(#) With a minimum of 5 degrees-of-freedom per node (2D flow velocity, pressure, density, salinity), a typical calculation has 720 x 720 x 20 x 5 ≈ 50 Million unknowns.

Definitions of V&V

- **Verification:** *"The process of determining that a computational model accurately represents the underlying mathematical model and its solution."*

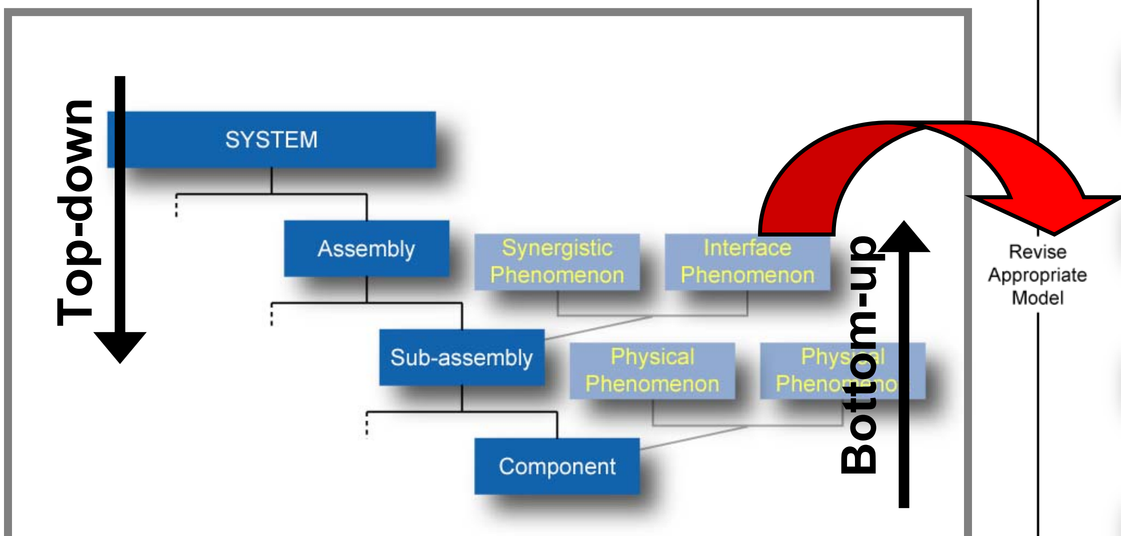
"Stability + Consistency → Convergence."
(Equivalence theorem of Peter Lax; Comm. in Pure and Applied Mathematics, 1954.)

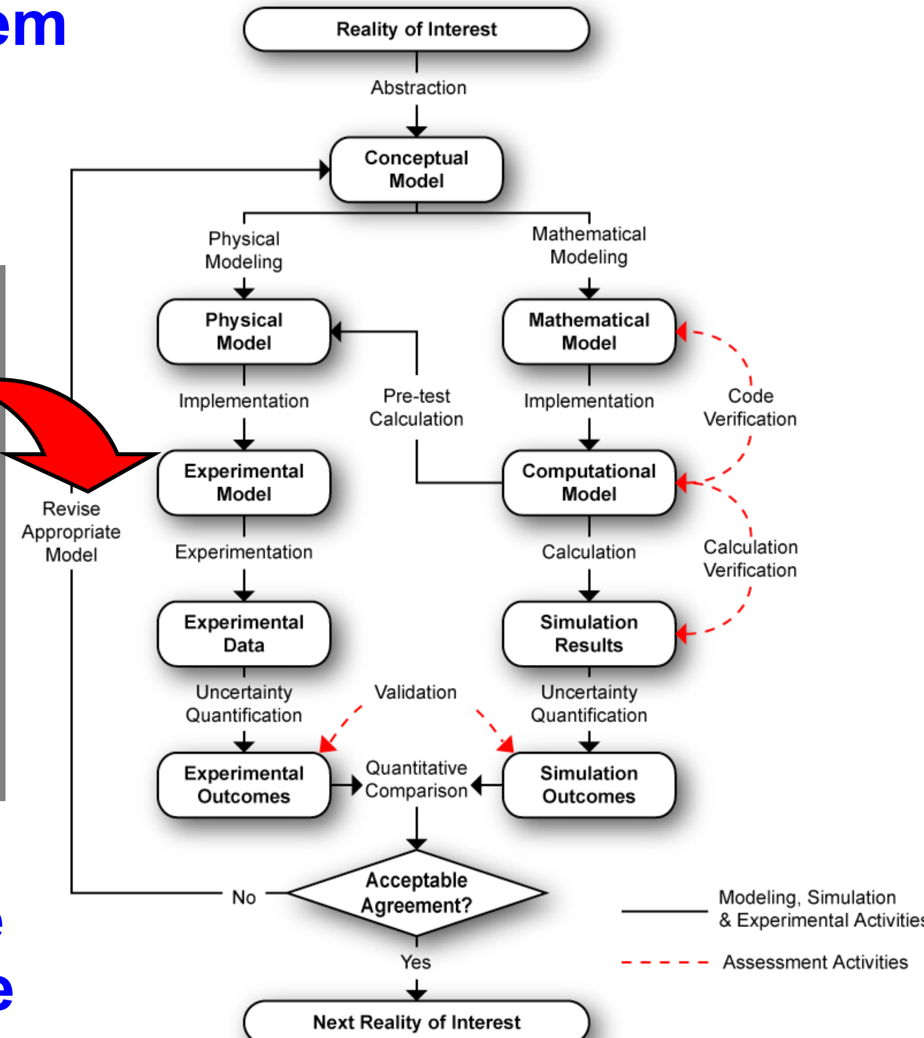
- **Validation:** *"The process of determining the degree to which a computer simulation is an accurate representation of the real world, from the perspective of the intended uses of the model."*

"Accuracy + Uncertainty + Robustness → Validation."
(My own definition, which matches common wisdom at U.S. DOE National Laboratories about what "validation" means.)

Typical Implementation of V&V

- The complexity of a system can be decomposed in a hierarchical manner.

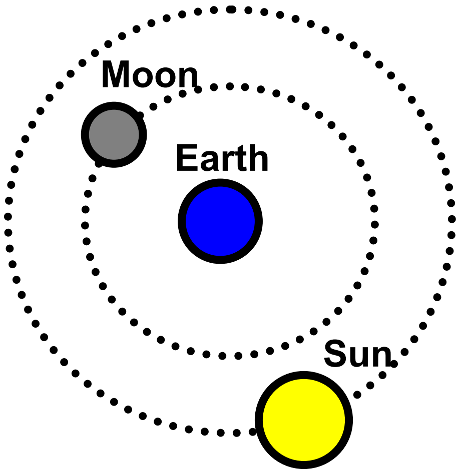




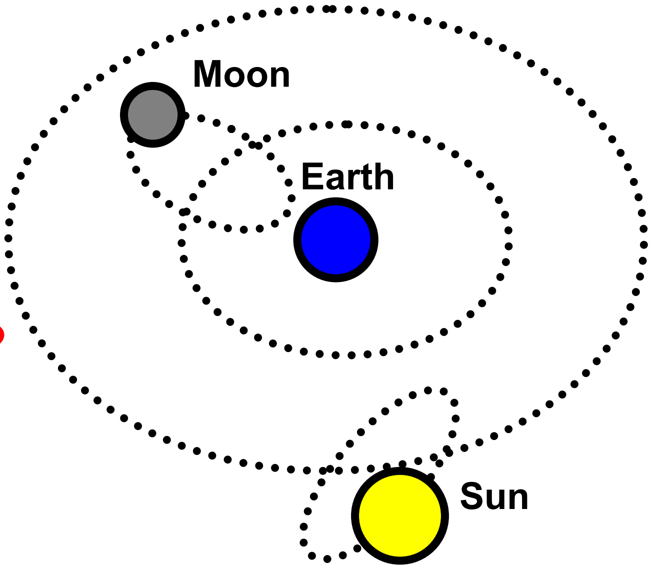
- V&V assessments can be performed either from the top-down or bottom-up.

The Example of Greek Astronomy

- These models dominated Western astronomy for over 2,000 years, reproducing the observation of planet positions and predicting the cycles of seasons with remarkable accuracy.



Are these models valid?



— Pythagoras & Aristotle (~500 BC)

— Ptolemy (~300 BC)

- According to the definition, the answer is *"yes."* They are validated for their intended purpose, which was to predict the cycle of seasons for growing crops.

The Points Made by This Example

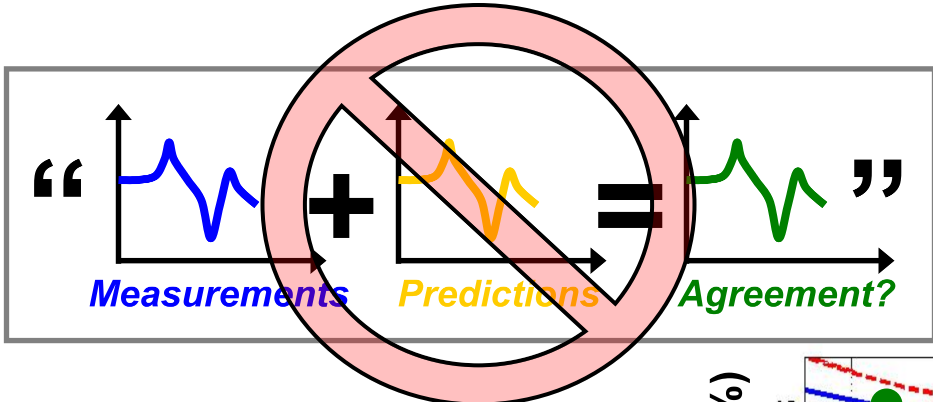
- The point of this (somewhat provocative) example is that model validation is not necessarily about *truth*; it is about *control*.
- One can always *"dial-in"* a desired level of accuracy through *"knob tuning"* or parameter calibration … But matching the test data does not necessarily uncover how the real world behaves.
- I argue that what is equally important to matching the test data is to control the modeling assumptions. One needs to understand which assumption is appropriate in what context.

V&V Should Support Decision-making

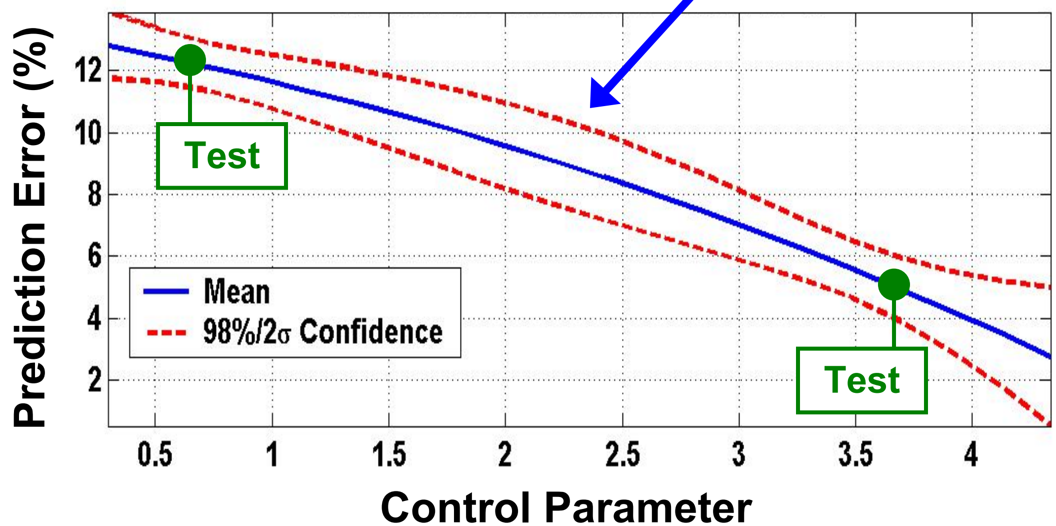
- The starting point of any V&V assessment should be an understanding of the decision that is supported by numerical simulations and physical experiments.
- One should start by discussing it with the decision-makers and stakeholders (such as code developers, analysts, engineers, customers) to understand what needs to be decided, what the requirements are, and define what is expected of V&V activities.
- The key point is that the decision made should not be *vulnerable* to assumptions upon which the numerical simulation is based. (This is what I call *"robustness."*)

What it Means to be *"Predictive"* …

- **A predictive capability is a code product for which the *accuracy*, *uncertainty*, and *robustness* of predictions are quantified (and found to be sufficient).**



… including away from settings that have been tested experimentally.(#)



(#) Caveat: As long as it is justified to *"extrapolate"* predictions of the code away from settings that have been tested experimentally.

Outline

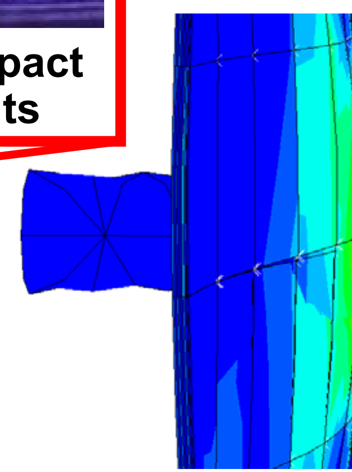
- Definitions and language of V&V
- An example in Structural Dynamics
- High-level description of a few V&V activities
- Two lessons (often learned the hard way)
- Overview of how V&V is structured at Los Alamos
- Closing remarks

Context of this Application

- This application is the impact of composite plates with a projectile to understand how the material fractures and how damage grows under cyclic loading.

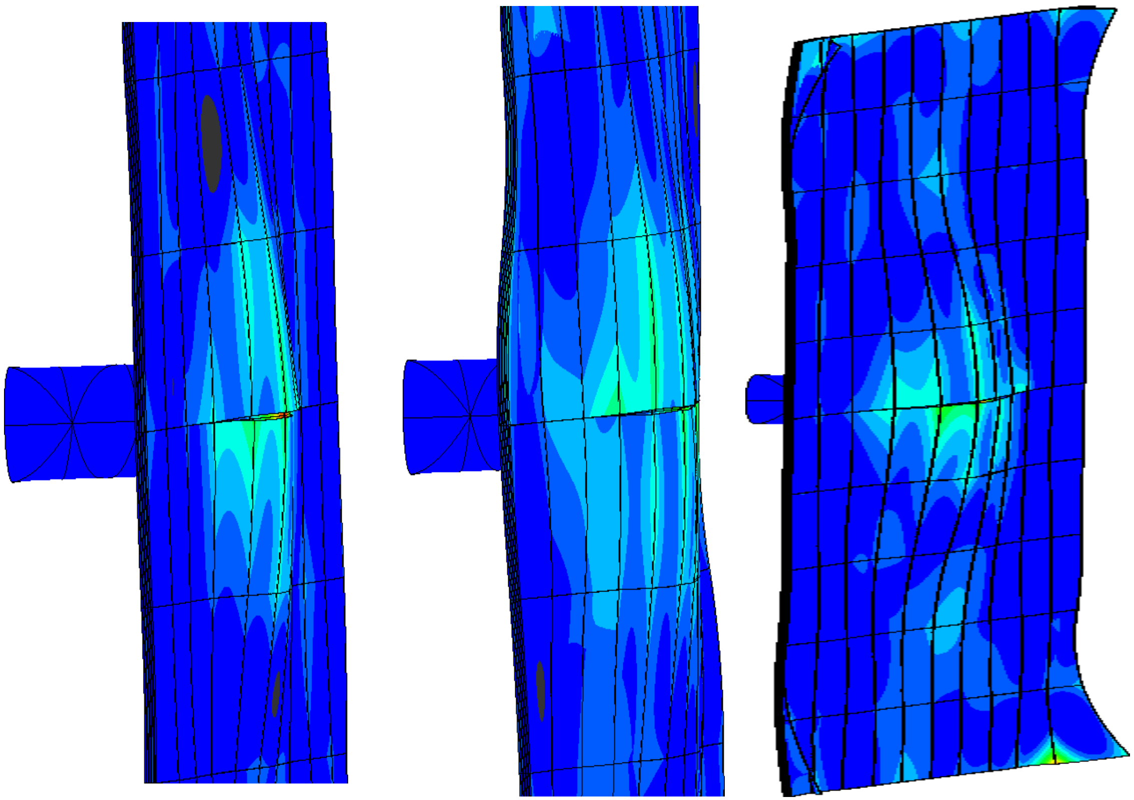
(Predator UAV pictures, courtesy of General Atomics.)

Projectile Impact Experiments

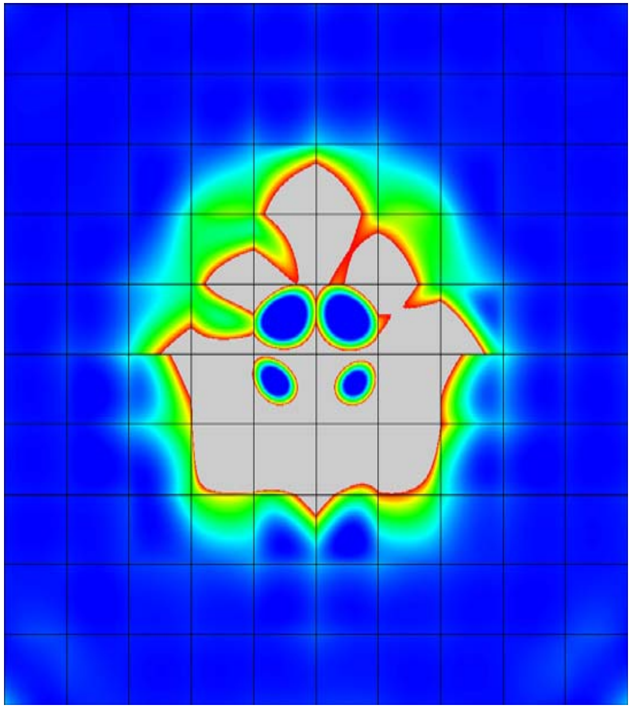


Impact Simulations

Example of Impact Simulation



Snapshots of Plate Deflection Due to Projectile Impact

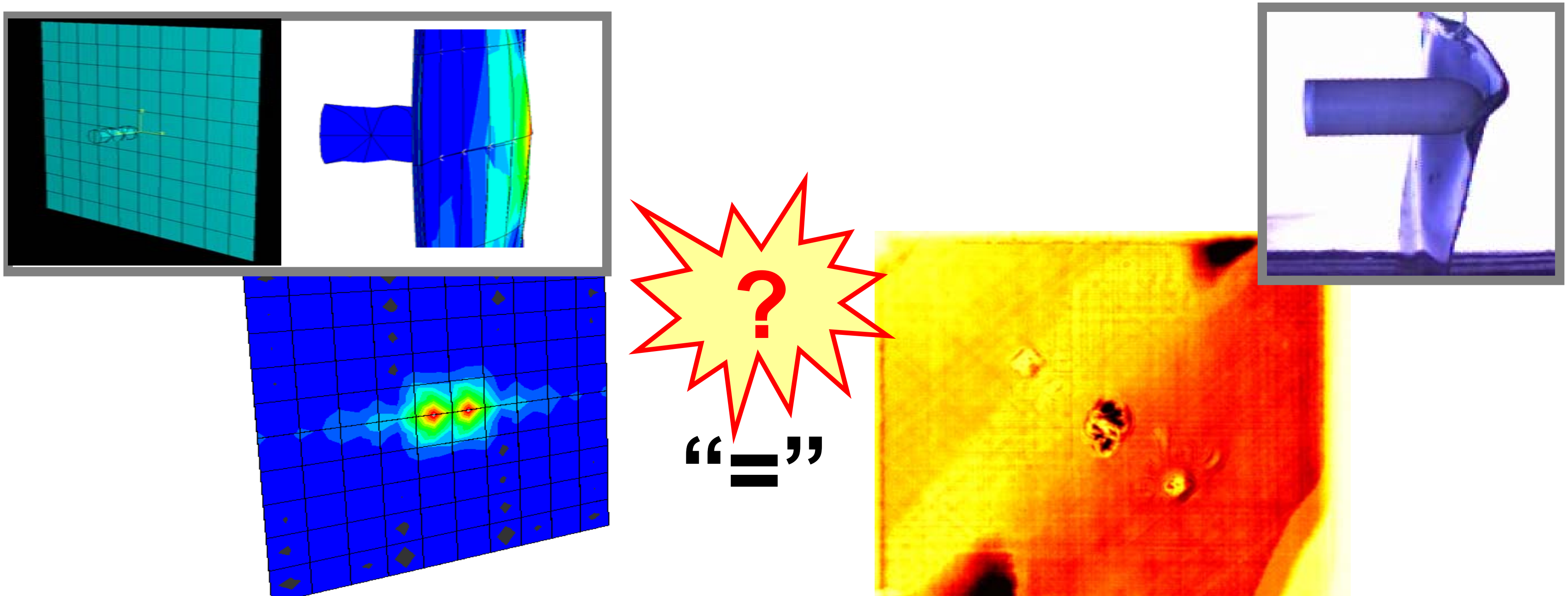


Prediction of Delamination

(Reference: LA-UR-05-6540. Credit: T. Tippetts.)

Modeling and Simulation (M&S)

- **The composite model, integrated to the finite element code, is capable of predicting the onset and growth of fiber splitting, delamination, and fiber fracture.**



Prediction of Delamination **Measurement of Delamination**

- **How *credible* are these predictions?**

Sources of Modeling Uncertainty

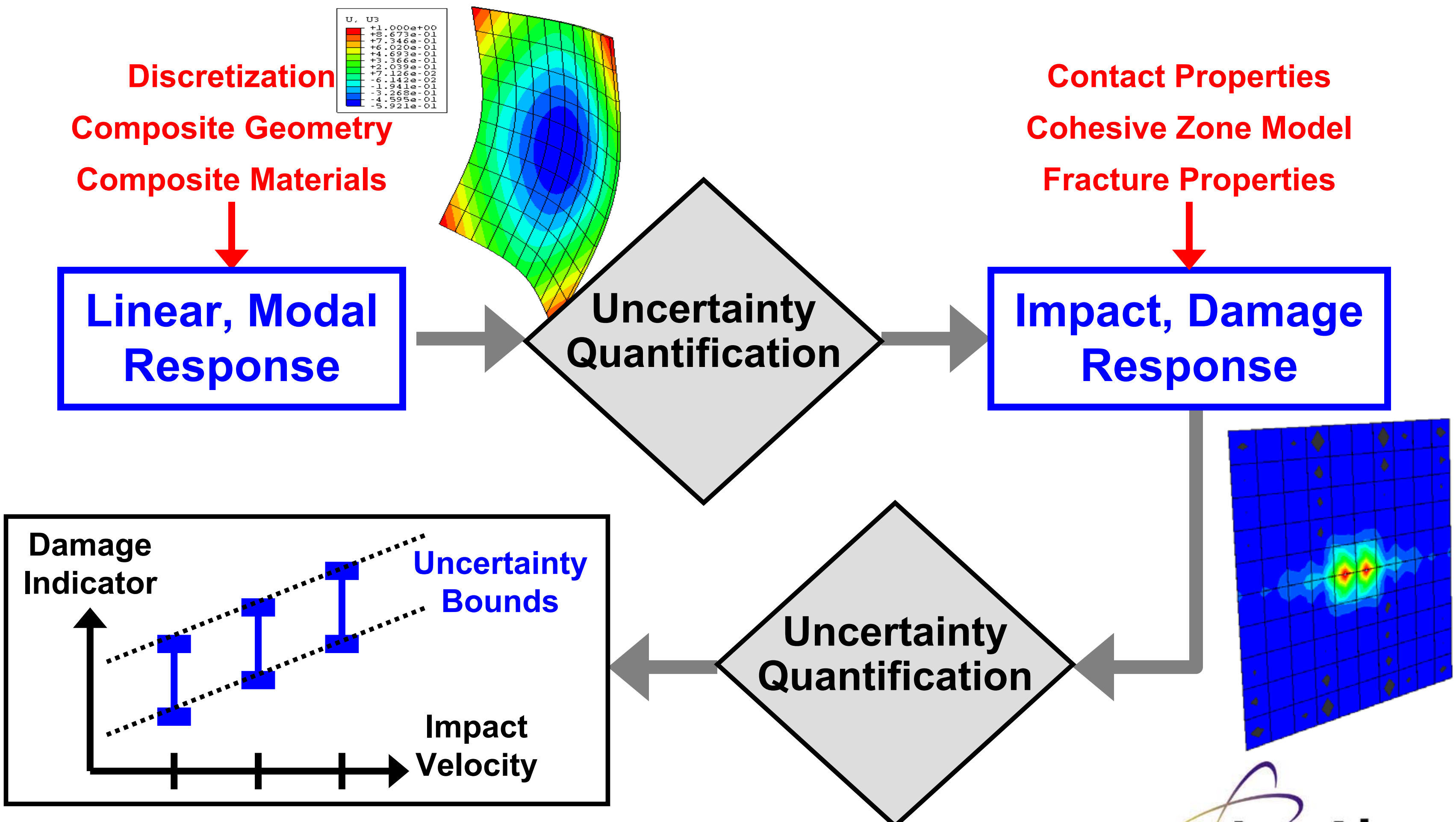
- **The simulation relies on many assumptions such as those embodied by the model of contact between the projectile and plate, or the model of material damage.**

Source	Corresponding Parameterization	Type
Ply orientation angles	θ_1, θ_2, θ_3, θ_4, θ_5, θ_6, θ_7, θ_8	*Variability*
Composite material	E_{11}, E_{22}, E_{33}, G_{12}, G_{13}, G_{23}, E_{12}, E_{13}, E_{23}, ρ	*Variability*
Fracture properties	$(G_C^{(k)}; T_{Max}^{(k)})$ for each fracture mode (k=I, II, III)	*Variability*
Impact velocity	V_I	*Variability*
Prediction of fracture	Fracture parameter, $\lambda_F \in [0; 1]$	*Fuzziness*
Cohesive Zone Model	Location of CZM finite elements	*Assumption*
Cohesive Zone Model	Shape of the force-displacement curve	*Assumption*
Projectile contact	Coefficient of the Hertz contact model, k_{NL}	*Assumption*

- **Which one controls the overall prediction uncertainty?**

Separate Effect Validation Tests

- **The first step is to validate the modal, linear response. The second step is to validate the impact response.**



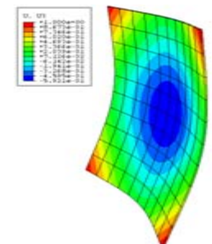
Validation of the Modal Response

- Simply comparing measurements to predictions may not indicate where the modeling error comes from.

ω_{Test} = 274.1 ± 2.9 Hertz

?

"="



ω_{Model} = 281.2 ± 19.8 Hertz

- V&V activities can guide the understanding of *where* the prediction error and uncertainty come from.

Source of Modeling Error or Uncertainty	Prediction Bias (μ)	Prediction Uncertainty (σ)
Solution Convergence	2.07 Hertz	0.00 Hertz
Reduced-order Modeling Error	0.73 Hertz	0.05 Hertz
Parametric Variability	0.00 Hertz	19.00 Hertz
Modeling Error	4.30 Hertz	4.73 Hertz

What is Learned?

- **What is learned, beyond quantifying the accuracy and uncertainty bounds of numerical predictions?**
- **… Which sub-models, parameters, or assumptions are controlling the overall prediction uncertainty.**
- **... Whether the modeling assumptions are appropriate (or not); when assumptions start to break down.**
- **... If additional physical testing is needed to improve the validation; what kind of tests, how many tests?**
- **... Whether the use of calibration is justified (or not).**
- **… What the extrapolation *"power"* of the model is.**

Outline

- Definitions and language of V&V
- An example in Structural Dynamics
- **High-level description of a few V&V activities**
- Two lessons (often learned the hard way)
- Overview of how V&V is structured at Los Alamos
- Closing remarks

Objectives of V&V

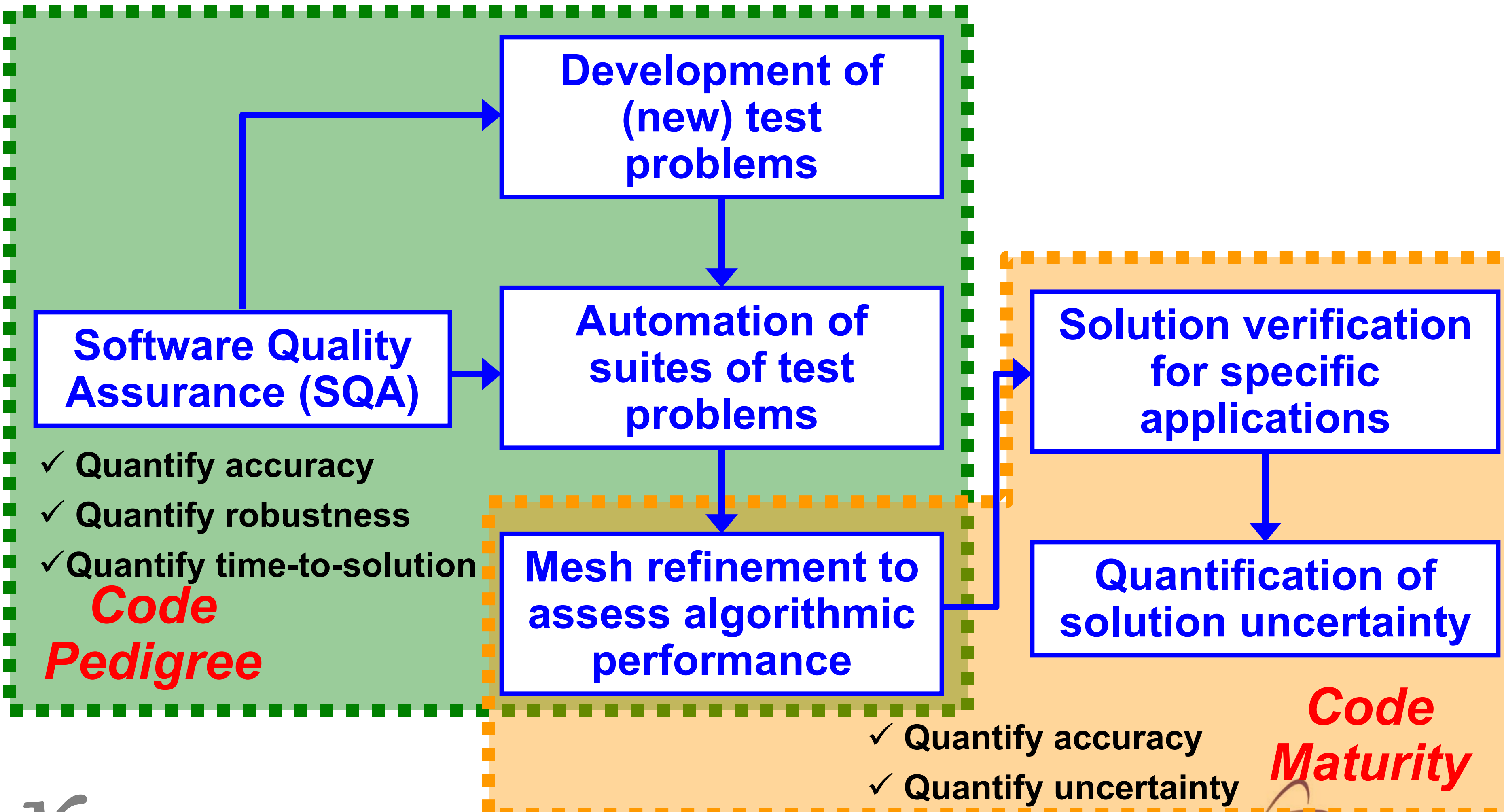
- Build *"pedigrees"* for code products and assess their maturity for application to specific problems.
- Quantify the potential sources of uncertainty and lack-of-knowledge, and their effects on predictions.
- Support decision-making by quantifying uncertainty. (The prediction uncertainty must be estimated to calculate a probability of success or failure, reliability, risk, or margin-to-requirement.)
- V&V activities can also help to establish priorities for experimental campaigns, model implementation, and code development.

Typical V&V Activities

- Code verification activities
- Response feature extraction
- Asymptotic convergence of discrete solutions
- Local sensitivity study (finite difference-based)
- Design of computer experiments
- Global sensitivity (variance-based), effect screening
- Development of fast-running meta-models
- Uncertainty propagation and assessment
- Test-analysis comparison and correlation
- Model revision and parameter calibration
- Extrapolation of prediction accuracy and uncertainty

Possible Feed-back

Verification activities map directly into assessing code *pedigree* and *maturity*.



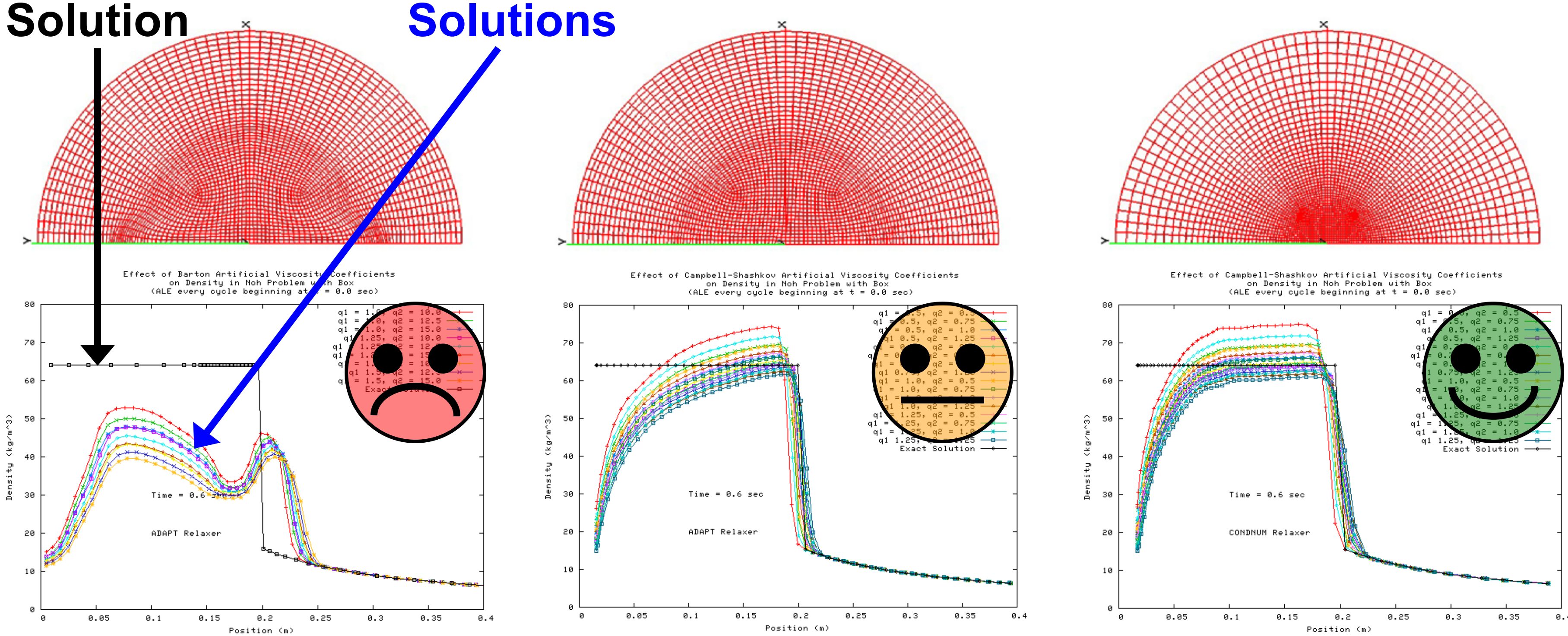
Code Verification Test Problems

- **Running test problems to compare exact and discrete solutions is the main *“workhorse”* of code verification.**

Example of a converging flow of material:

Exact Solution

Discrete Solutions

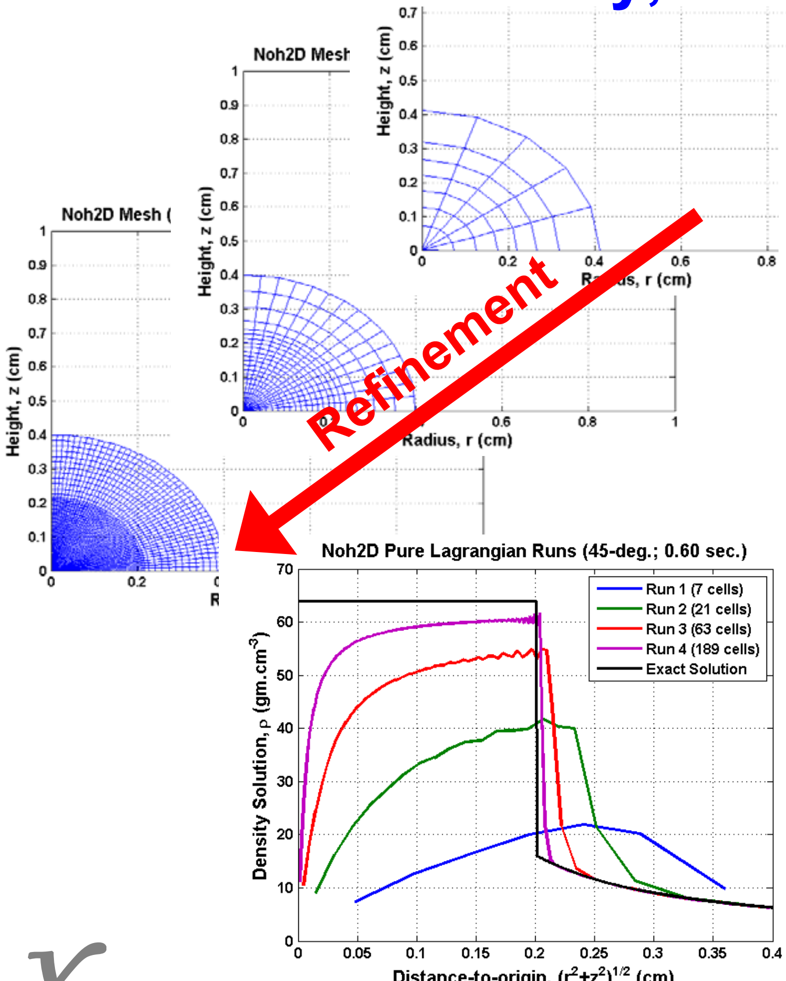


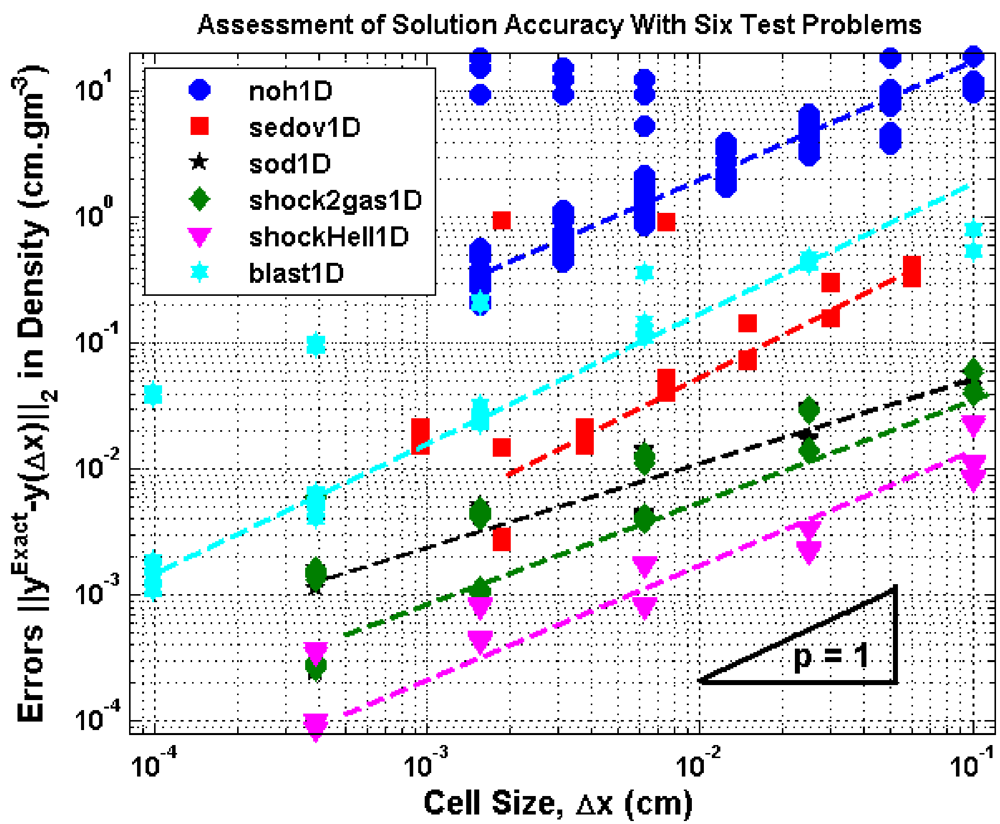
Algorithm A **Algorithm B** **Algorithm C**

(Reference: LA-UR-08-6011.)

Mesh Refinement Studies

- Mesh refinement studies assess algorithmic robustness, solution accuracy, and quantify solution uncertainty.





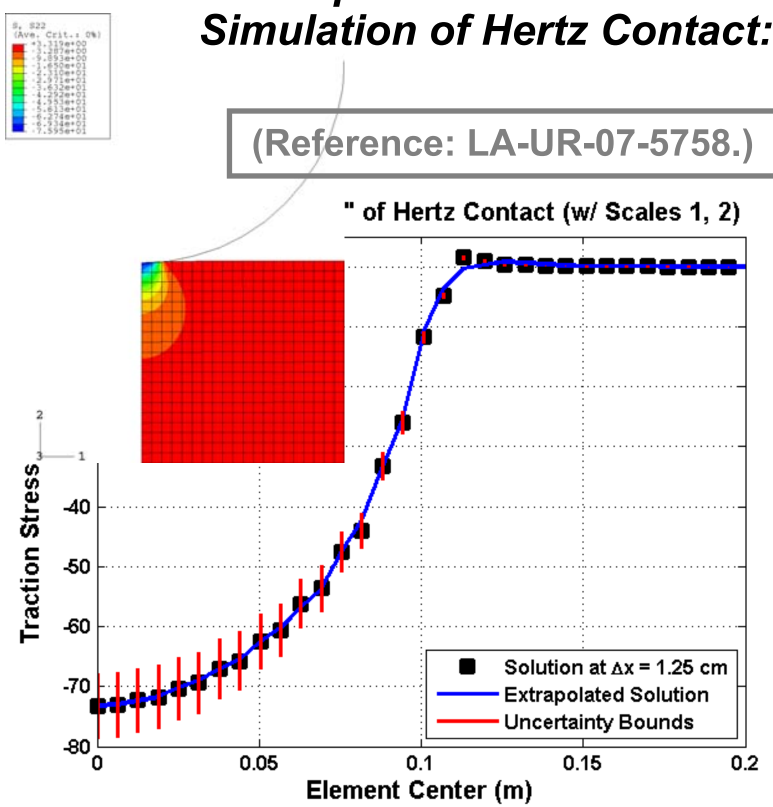
(Reference: LA-UR-07-7768.)

Quantification of Solution Uncertainty

- The *solution uncertainty* caused by running a calculation with a given level of resolution *"Δx"* must be quantified because it contributes to the overall uncertainty budget.

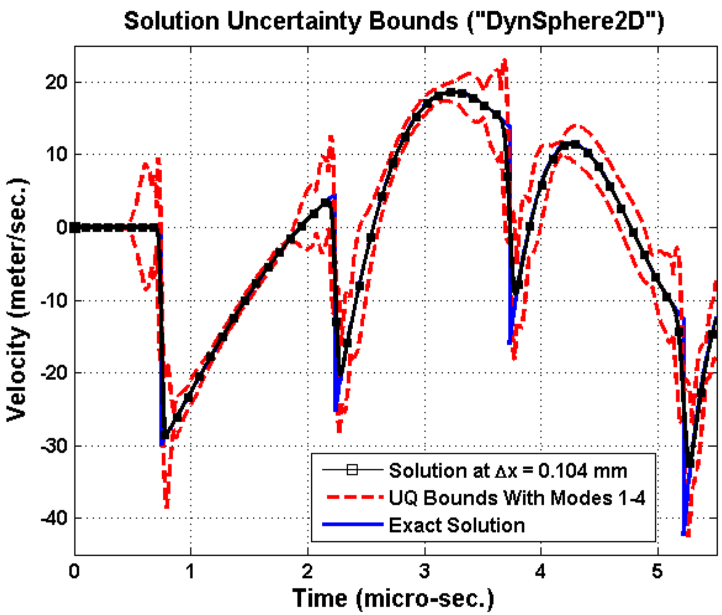
Example of Finite Element Simulation of Hertz Contact:

(Reference: LA-UR-07-5758.)

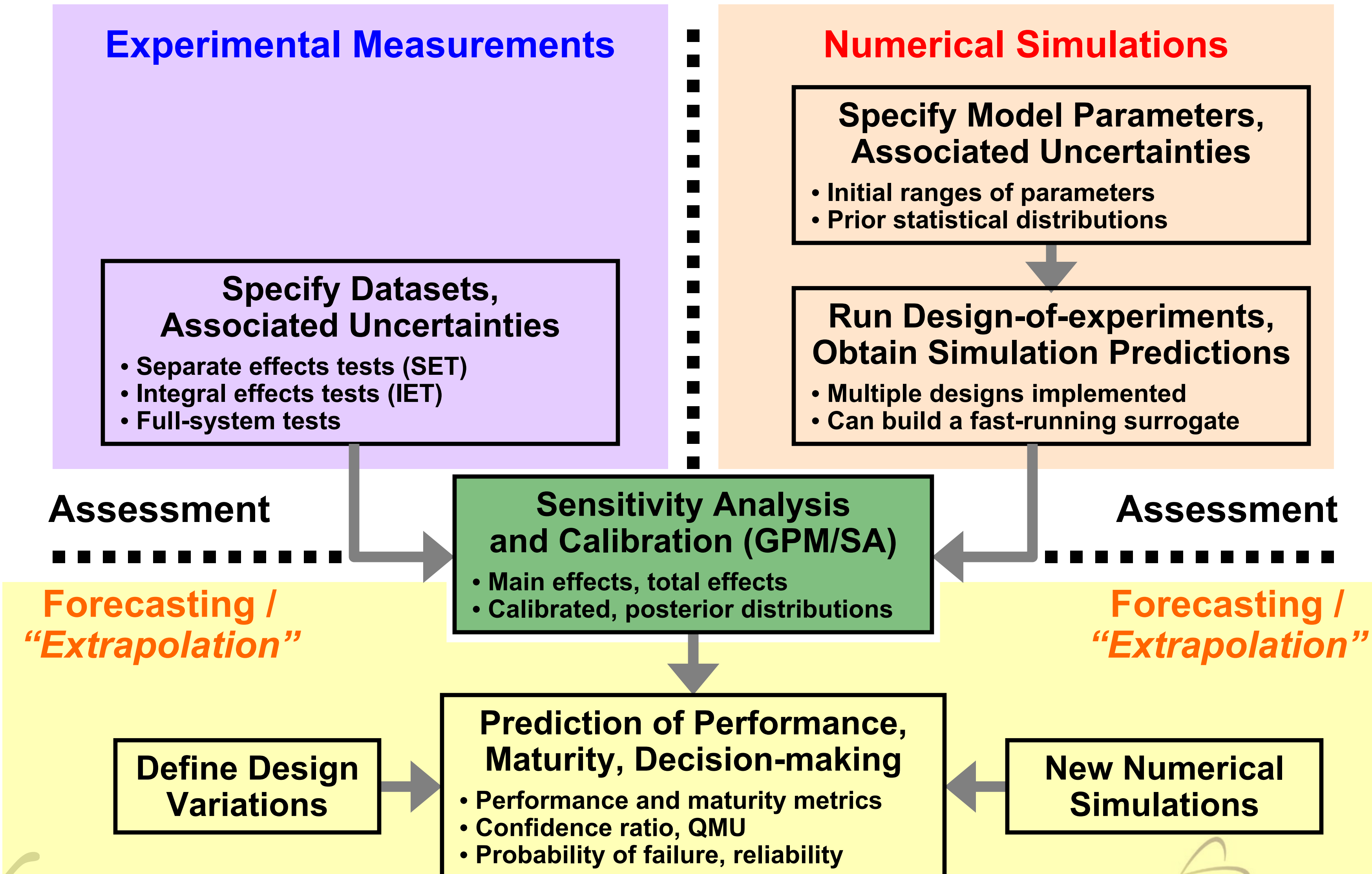


Example of Hydro-dynamics Simulation of Reflecting Waves:

(Reference: LA-UR-08-4197.)

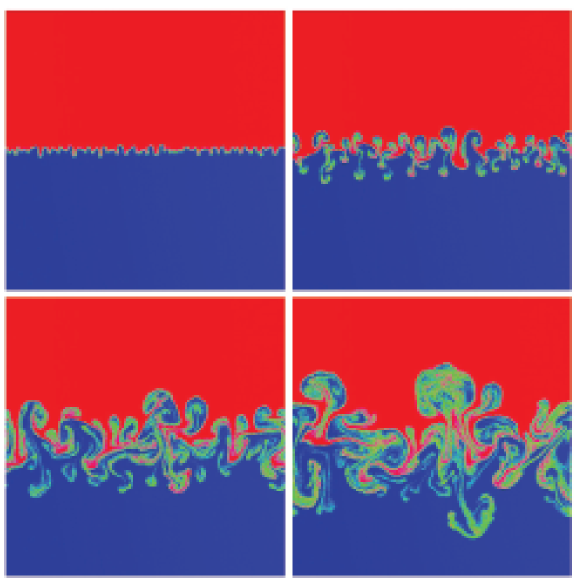


Uncertainty Quantification (UQ)



Sensitivity Analysis (SA)

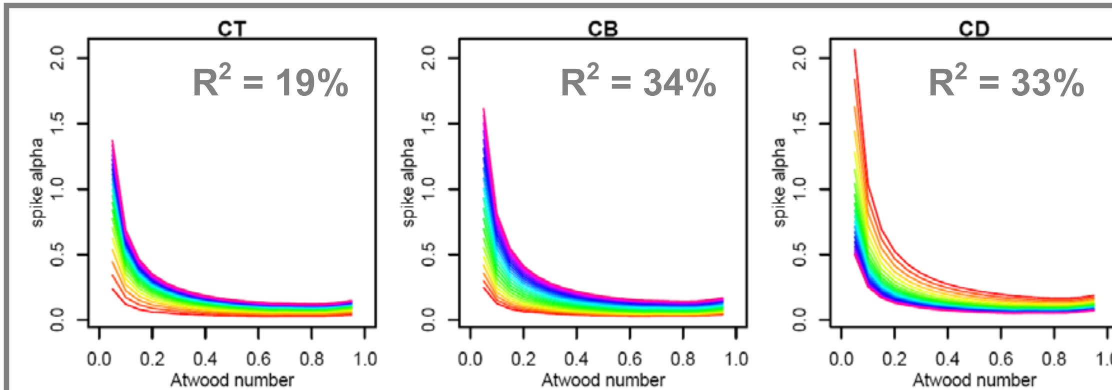
- **The statistical influence that input parameters, or effects, have on output predictions is assessed to *guide* model development and the design of experimental campaigns.**



Simulation of Rayleigh-Taylor Instability

Flow Regimes, Initial Conditions, Materials, etc. → **Numerical Model:** $\frac{\partial y}{\partial t} + \nabla F(y;t) = S(t)$ → **Instability Growth Rate, L = $\alpha A g t^2$**

Numerical Model ⇅ **Sub-grid Closure Model:** $p = F(\rho;e;C_B;C_T;C_D)$

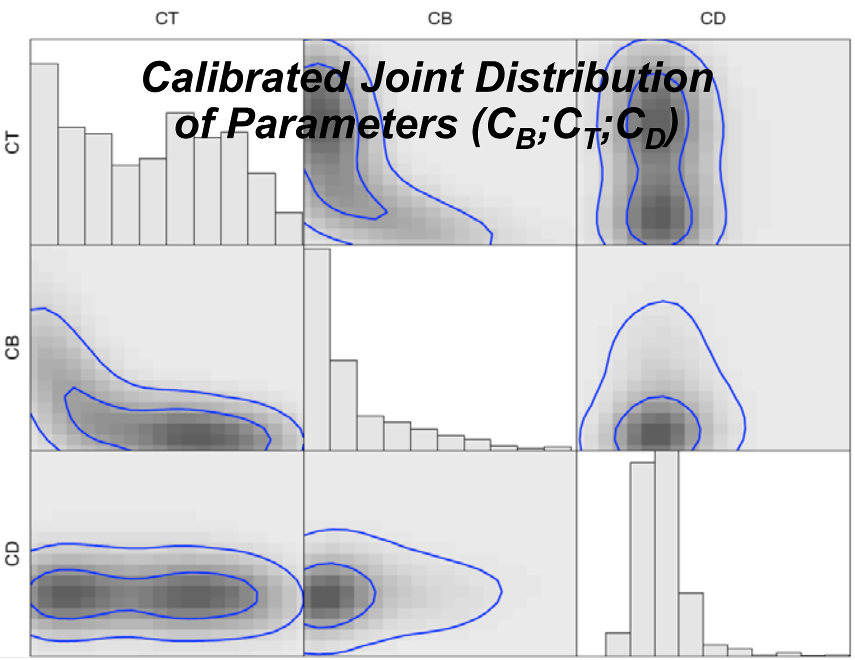


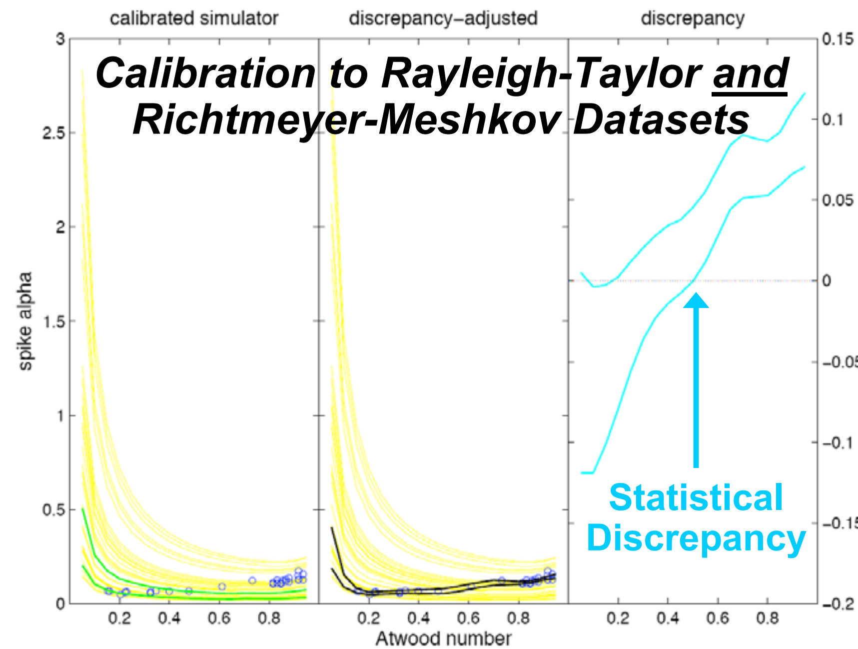
Sensitivity Analysis of the k-L Turbulent Mix Model

ASC

Parameter Calibration

- **Calibration of model parameters can be performed, not so much to match the experimental observations, but to *infer* modeling uncertainty from experimental variability.**





- **Techniques for statistical calibration can estimate any unresolved *discrepancy* that the calibration process is unsuccessful compensating for.**

Outline

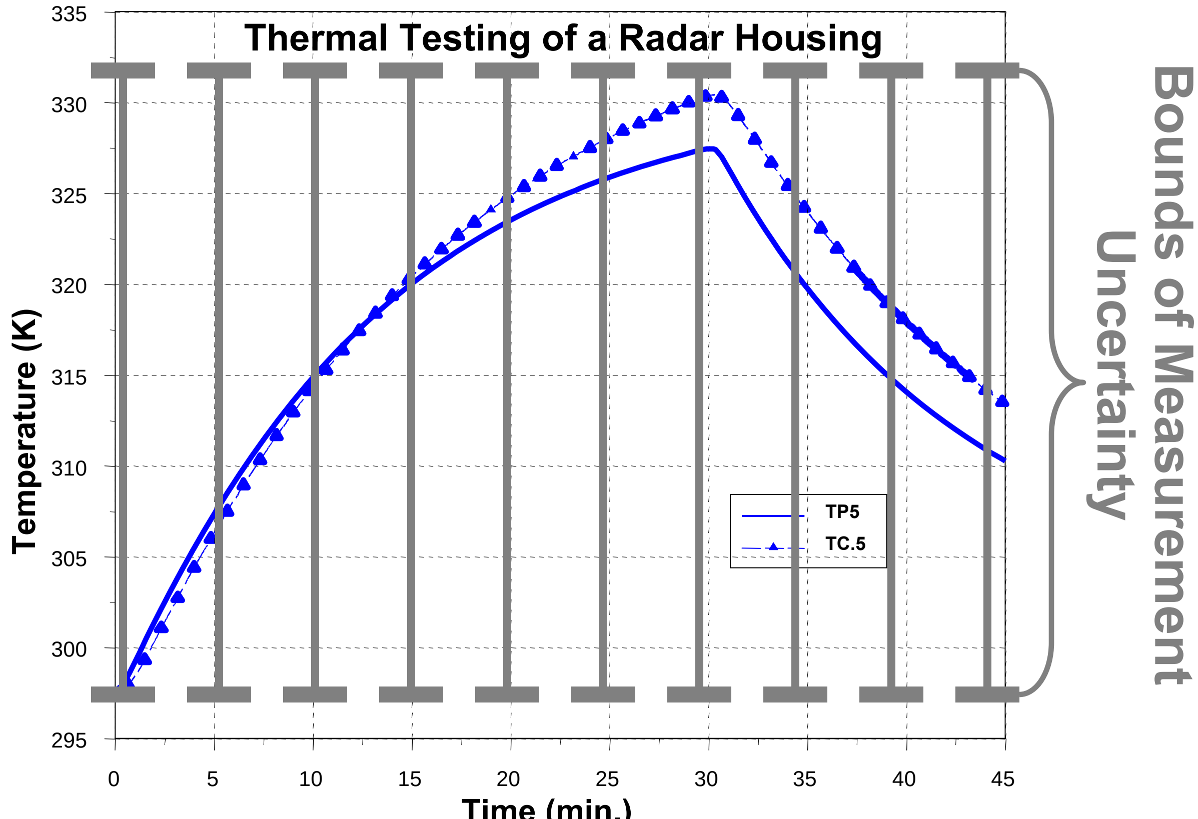
- Definitions and language of V&V
- An example in Structural Dynamics
- High-level description of a few V&V activities
- **Two lessons (often learned the hard way)**
- Overview of how V&V is structured at Los Alamos
- Closing remarks

First Lesson

- **Measurements without experimental error bounds are meaningless. (Replicates, replicates, replicates …)**

Tim Trucano

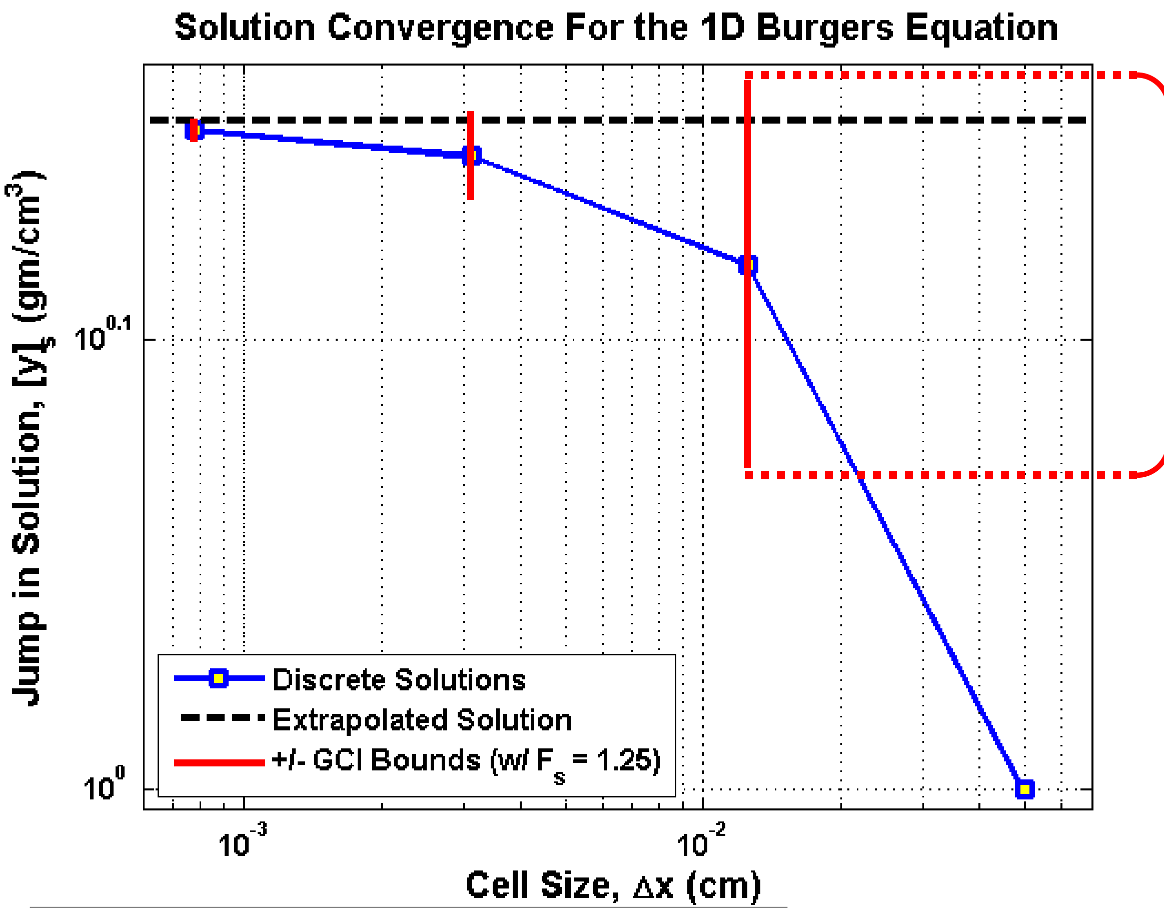
V&V Pioneer at Sandia National Laboratories, New Mexico



"If your measurement error is not quantified, then it must at least be the size of your graph." **(Reference: T. Trucano, NECDC|06 Conference, Los Alamos, NM, October 2006.)**

Second Lesson

- **Predictions without numerical uncertainty bounds are meaningless. (Do your mesh-refinement homework!)**



1D Burgers Equation

$$\begin{cases} \dfrac{\partial y(x;t)}{\partial t} + \dfrac{\partial F(x;t)}{\partial x} = 0 \\ F(x;t) = \dfrac{1}{2} y^2(x;t) \\ -\dfrac{L}{2} \le x \le \dfrac{L}{2}, \quad 0 \le t \le 1 \\ y(x;t=0) = y_o(x) \end{cases}$$

(Picture extracted from LA-UR-06-8078.)

Outline

- Definitions and language of V&V
- An example in Structural Dynamics
- High-level description of a few V&V activities
- Two lessons (often learned the hard way)
- **Overview of how V&V is structured at Los Alamos**
- Closing remarks

Outline

- Definitions and language of V&V
- An example in Structural Dynamics
- High-level description of a few V&V activities
- Two lessons (often learned the hard way)
- Overview of how V&V is structured at Los Alamos
- Closing remarks

… Repeat of Main Points

- Verification and Validation (V&V) is a rigorous and scientifically defendable procedure by which analysts can demonstrate the *credibility* of their predictions.
- The overarching goal is to identify *all significant* sources of uncertainty and lack-of-knowledge, and *quantify* the effects that they have on predictions.
- If an assumption cannot be justified while, at the same time, exercising a significant influence on predictions, then results are vulnerable to a lack-of-knowledge …
- … And poor *robustness* should be improved by learning the *"unknown"* physics through, for example, small-scale experiments. This is as important as the *accuracy* of predictions.

Don’t Turn to the “Dark Side”