Approved for Unlimited Release
LA-UR-07-2575

Overview of the LANL Verification and Validation Program

Jerry S. Brock, Ph.D.

Code & Calculation Verification Project Leader (jsbrock@lanl.gov)

Scott W. Doebling, Ph.D.

Verification & Validation Program Manager (doebling@lanl.gov)

Salishan Conference on High-Speed Computing - April 2007

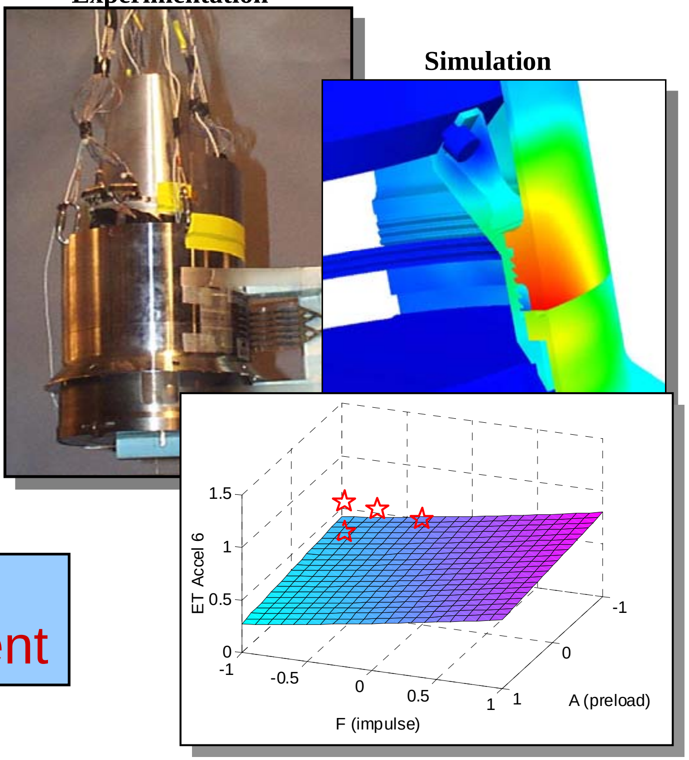
LANL Verification & Validation (V&V) Program Overview: Outline

- Advanced Simulation & Computing (ASC) Program
- Motivation for Verification & Validation
- The V&V Program at Los Alamos
- The V&V Process
  - – Software Life Cycle
  - – Hierarchical V&V Assessment
- Example V&V Assessments
- Present V&V Challenges
- V&V Computing Requirements
  -- Capability and Capacity

V&V: Quantitative Bridge Between Physics Models, Simulation & Experiment

Experimentation

Simulation



Quantitative Comparison

Advanced Simulation & Computing Program at Los Alamos National Laboratory

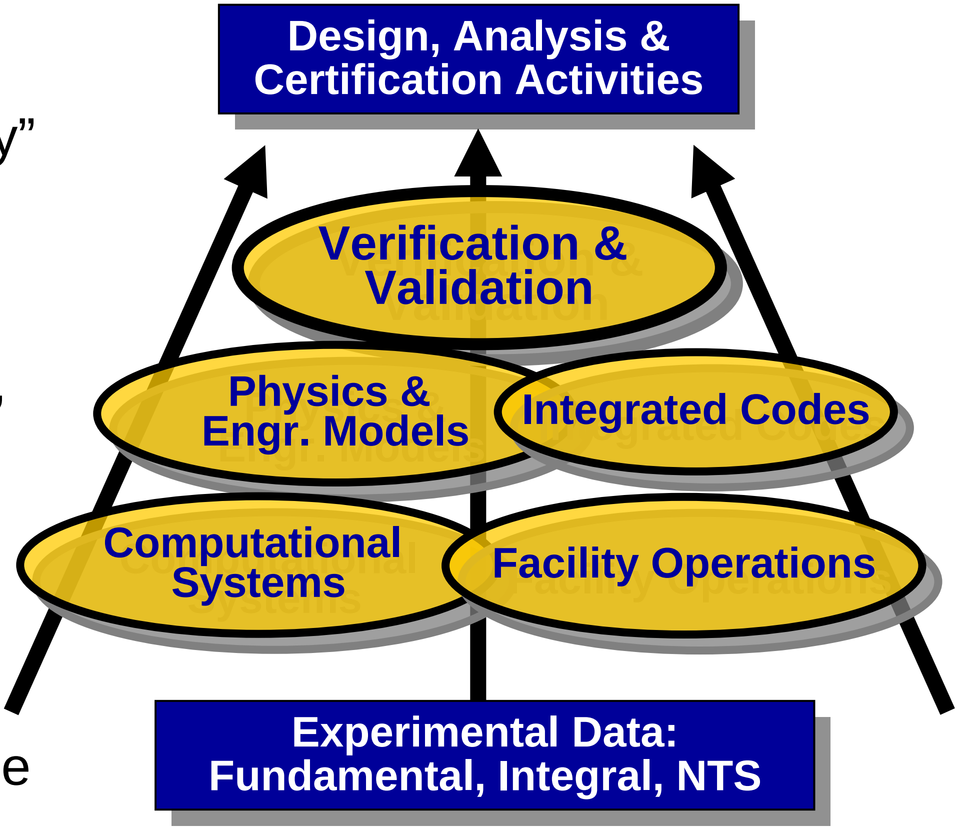
Dr. John Hopson, Program Director
Dr. Ralph Nelson, Deputy Program Director
Dr. Cheryl Wampler, Deputy Program Director

- **Physics & Engineering Models** — **Dr. Paul Maudlin**
- **Integrated Codes** — **Dr. Marvin Alme**
- **Verification & Validation** — **Dr. Scott Doebling**
- **Computational Systems & Software Environment** — **Dr. John Thorp**
- **Facility Operations & User Support** — **Dr. Cheryl Wampler**

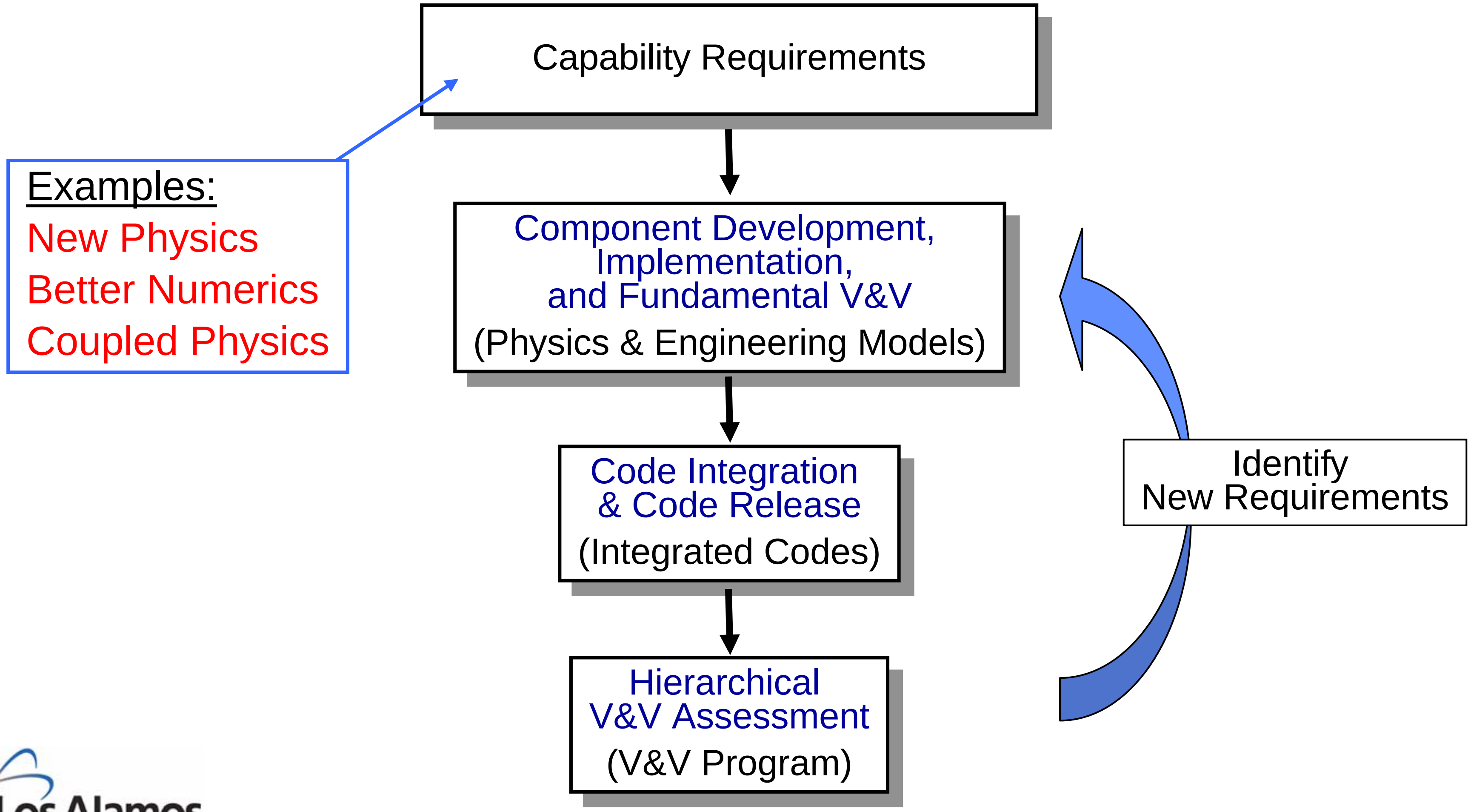
ASC: Develop, demonstrate, assess and deploy simulation tools to "analyze and predict the performance, safety and reliability of the nations nuclear weapons and to certify their functionality"

Verification & Validation Unites Basis for Calculation Credibility

- Verification
  -- "Solving the equations correctly"
  – i.e. get the math right
- Validation
  – "Solving the correct equations"
  – i.e. get the physics right
- Defensible Assessments
- Application-Oriented Focus
- Hierarchical Assembly of Evidence
- Error & Uncertainty Quantification



ASC Capability Development Process



What Makes Our Validation Tasks Difficult?

Simulation Issues

- Complex and often coupled physical phenomena
- Simulation of nuclear diagnostics
- Complex geometries with 3-D characteristics
- Sensitivity to temporal and spatial resolution
- Computational requirements for all of the above

Physical Environment Issues

- High temperatures, pressures and energy densities
- Nuclear reactions
- Extremely short event durations

Experimental Data Issues

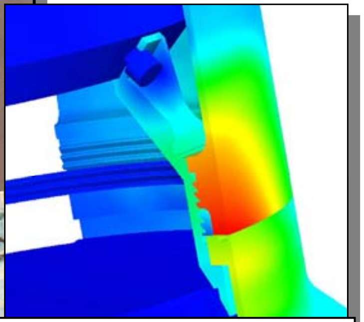
- Fundamental data in regimes of interest often difficult to obtain
- Little data on test-to-test repeatability and unit-to-unit variability
- Reliance on historical data for many environments

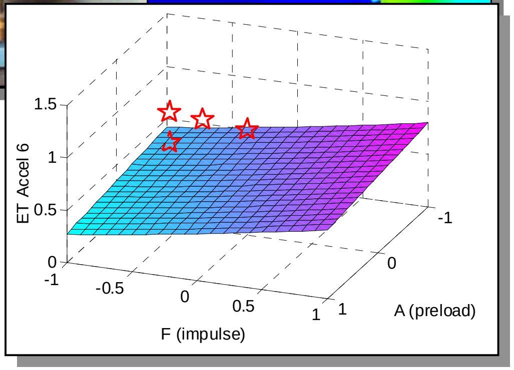
LANL Verification & Validation Program

- V&V Motivation
- V&V for Predictive Capability
- Program Structure / Roles & Responsibilities
- Assessment Process & Hierarchical Approach
- Practical Issues & Improvements to V&V Process / Program Integration

Experimentation

Simulation





Quantitative Comparison

Verification & Validation Motivation

- **Mission** - leadership in the development, implementation, and improvement of scientifically rigorous methods for the **assembly of evidence** to determine and document:
  - The **scientific credibility** of calculations from ASC codes, and
  - The **scientific utility** of these calculations for programmatic mission

- **Scientific credibility** means understanding the degree to which the codes produce the **right answer for the right reason**

- **Scientific utility** means ensuring that the code can address the **right questions** regarding complex scientific phenomena

Confidence in simulation extrapolation comes via confidence in physics & numerics models, *not calibration to experimental data*

What is Predictive Capability?

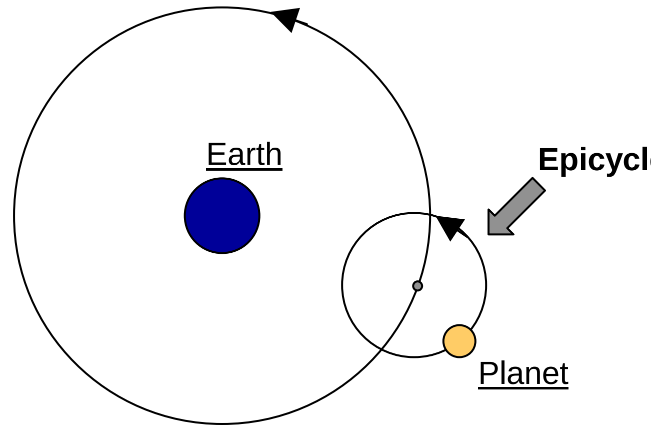
- Predictive science is about getting:
  - Right answers (including errors & uncertainty)
  - To the right questions (applications oriented)
  - *For the right reasons* (physics-based models)
- Elements of predictive capability include:
  - Moving from credible interpolation to credible extrapolation
  - Reducing reliance on calibration with integral data
  - Maintaining credible estimates of *numerical error and numerical/physical uncertainty on calculation results*

Calculations that give the right answer after calibration can be very useful, but predictive capability is about getting that right answer for the right reasons

Usefulness of Calibrated Model: Ptolemy's Model of the Solar System

Claudius Ptolemaeus (circa 100AD)

- Geocentric model: planets orbit in *epicycles*, where the center of each epicycle orbits the earth; it was the accepted model for over 1300 years



- Very accurate predictions of planets' positions - As measurement accuracy improves, more epicycles can be added to improve model accuracy
- Explains retrograde motion of planets, why planets are brightest when moving retrograde
- Geocentricity was compatible with then-dominant philosophy

..BUT!

- No underlying physical basis – WHY do the planets behave in this manner?
- Additional epicycles significantly increased model complexity and eventually called in to question the fundamental assumptions
- Did not explain some observations (e.g. why Mercury and Venus stay close to Sun)

Lessons for Predictive Science

- Having "good agreement" between calculations and observations is **not sufficient to establish scientifically credible predictive capability**.
- Note that this model was still useful, but not predictive!

LANL V&V Program Structure

Independent from Code Development and Code Application

Verification & Validation Program

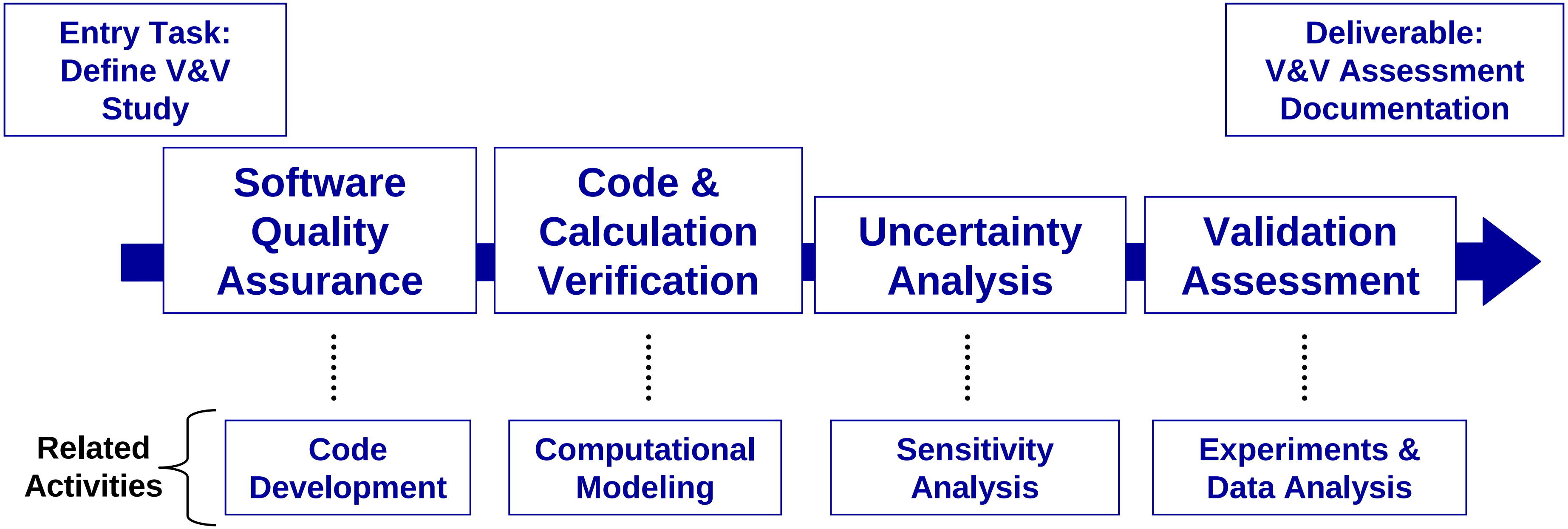
- **V&V Assessments for Applications of Interest**
- **Code and Calculation Verification: Test Problems & Numerical Error**
- **Analysis and Archiving of Experimental Data**
- **Uncertainty Analysis Methods and Applications**
- **Software Quality Assurance**

Why Independent V&V Program?

- Can make objective assessment of code behavior without a vested interest in the outcome
- Requires teams with diverse technical specialties working together

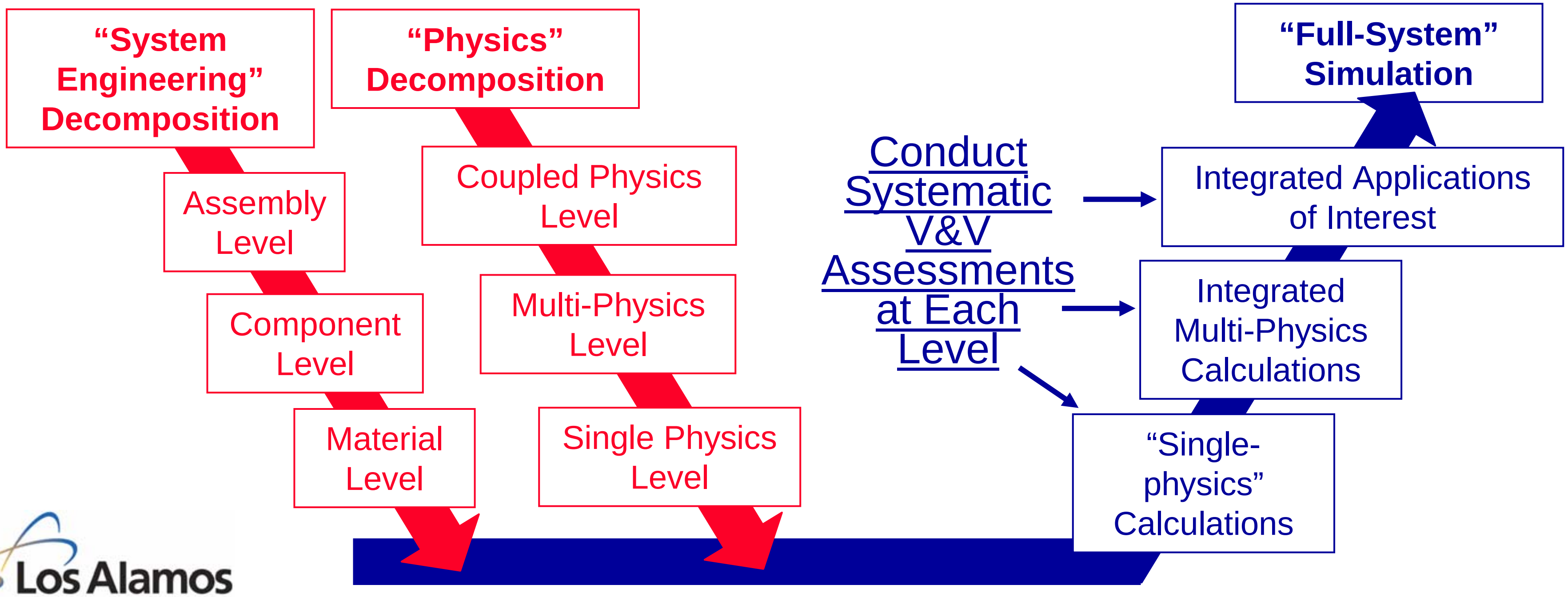
Basic Process for Systematic V&V Assessments

- Focus is on *assembly of evidence* via a systematic process
- Assessments can be applied at any level in the hierarchy



Hierarchical Approach to V&V Assessments

- **Hierarchical decomposition** of complex coupled-physics systems
- A systematic V&V assessment process is used to determine calculation credibility and utility at each level
- Credibility of system-level calculations is established via assessment of incrementally more complex calculations

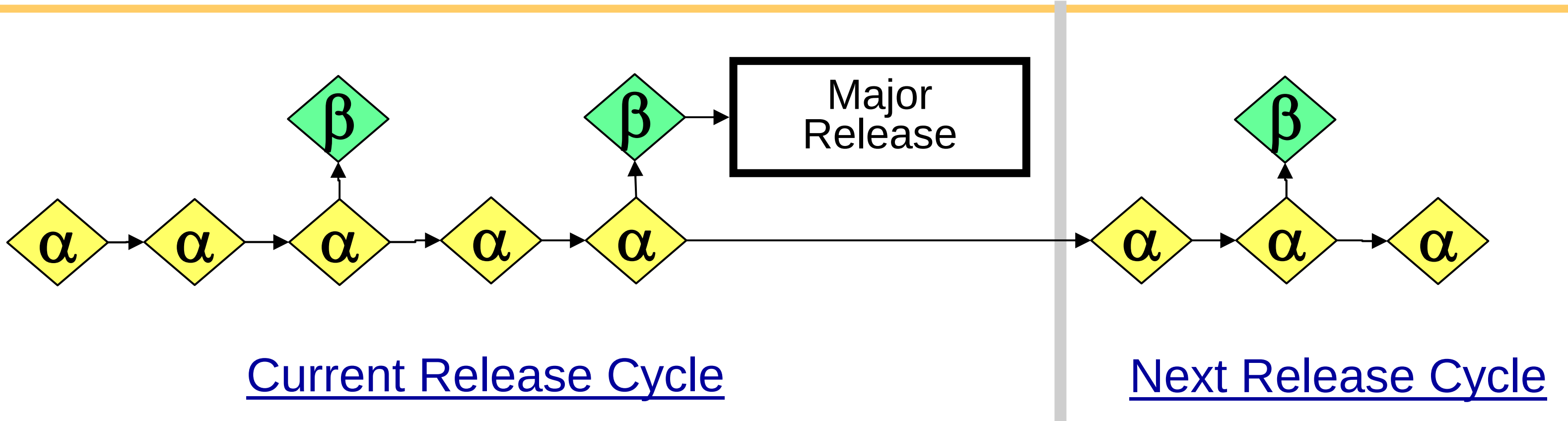


Practical Issues with V&V Assessments

- **Timing** Assessments do not exist for every code version
  - Code development is dynamic and ongoing during assessments
  - During assessment, additional code releases are made and adopted by some or all of the user community
  - Result: V&V assessments are not fully useful to end-users, i.e., assessment time scale does not allow investigation of each code version
- **Breadth** Specfic assessments focus on some physics and neglect others
- **Progress** No explicit measure of “improvement of predictive capability”
- **Sufficiency** No basis for determining “how much is enough?”
  - Trucano (NECDC 2006) notes that “sufficiency raises challenges of accumulation, communication and preservation of information”

How to improve the V&V Assessment Process ?

Code Version / V&V Assessment Challenge Solution is Based on Sufficient Code Pedigree



- Allow V&V assessment to exercise multiple alpha/beta releases
- Tighten loop between V&V and code team for faster impact of V&V assessments (e.g. weekly Crestone V&V meetings)
- Push verification and validation test cases upstream in the software testing cycle -- **"Born V&V'ed"**
- Pedigree of code generates confidence in the consistency of code between various versions – i.e. "near-neighbor" versions

V&V assessment is associated with a "release cycle" but not directly hard-wired to the "release version"

Born V&V'd: Closing the Loop Between Code Development & V&V Assessments

- **Link Code Development/Assessment** Capture the analysis of each V&V assessment in a reproducible manner, i.e., close the loop.
- **Solution** Make code acceptance tests an additional deliverable from the V&V assessments
  - These acceptance tests are then added to the code test library for future releases of the code
  - These tests should have clearly defined acceptance (pass/fail) criteria
- **Born V&V'd** refers to the demonstration of an inherent level of credibility for a code release via the passing of hierarchical acceptance tests
  - *This version of the code is Born V&V'd with respect to the following acceptance tests …………*

Improved Code Development/Assessment Process !

Practical Improvements for V&V Assessments

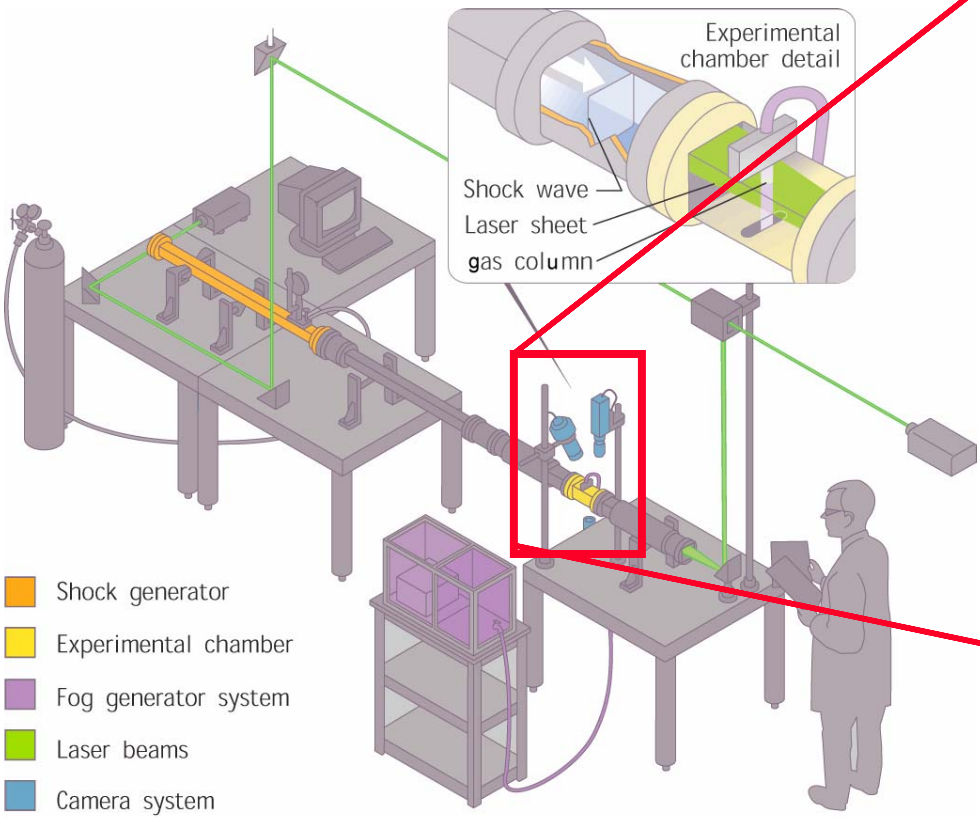
- **Timing** Assessments do not exist for every code version
  - *Each code release* has *passed a well-defined set of tests*
- **Breadth** Specfic assessments focus on some physics and neglect others
  - *Each code release* has *passed relevant physics tests*
- **Progress** No explicit measure of "improvement of predictive capability"
  - *Improvements in predictive capability* for a particular physics area can be tied to the accuracy with which the set of acceptance tests are calculated
  - Define/Track Test Suites for Verification, Primary and Secondary Validation
- **Sufficiency** No basis for determining "how much is enough?"
  - *Accumulation*: Important V&V Assessment analyses are captured into the library of code-release acceptance tests
  - *Communication*: Clear definition of what tests the code release has passed
  - *Preservation*: Acceptance test results are archived for each code release

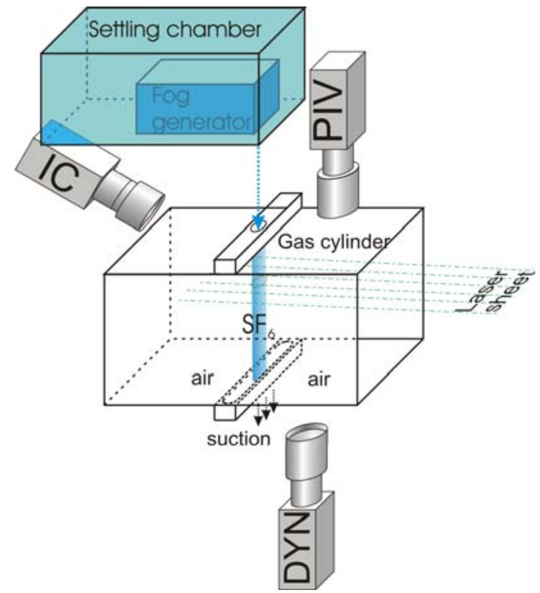
Example: Hierarchical Validation of Hydrodynamic Mixing in Stars

Increasing Complexity →

Hierarchical Level	Scenario	Reference Solution
Integral System (Application of Interest)	Simulation of Stellar Behavior	Observatory Data (if available)
"Single, More Complex Physics"	Simulation of Shocked Gas Mixing	Data from Laboratory Experiment
"Single Physics"	Simulation of Mix Test Problem	Exact solution (if available); Self-convergence analysis

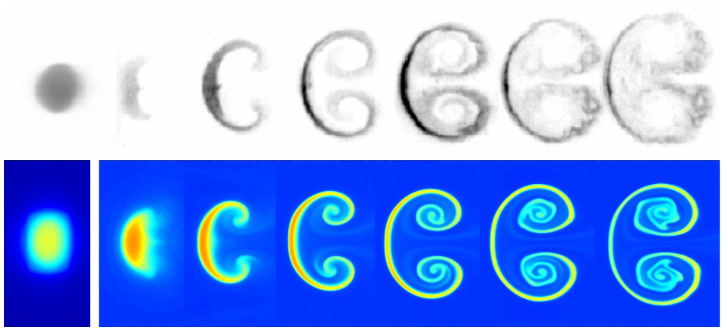
Hierarchical Validation: Validating Mix Models Using Gas Shock Tube Experiments





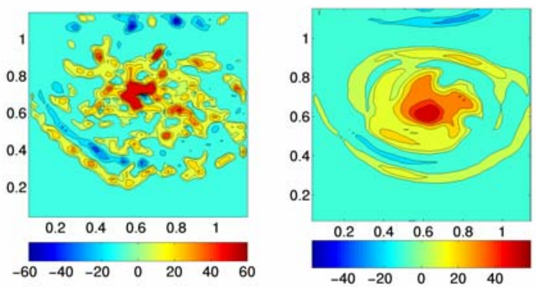
Test section with diagnostics

Hierarchical Validation: Validating Mix Models Using Gas Shock Tube Experiments

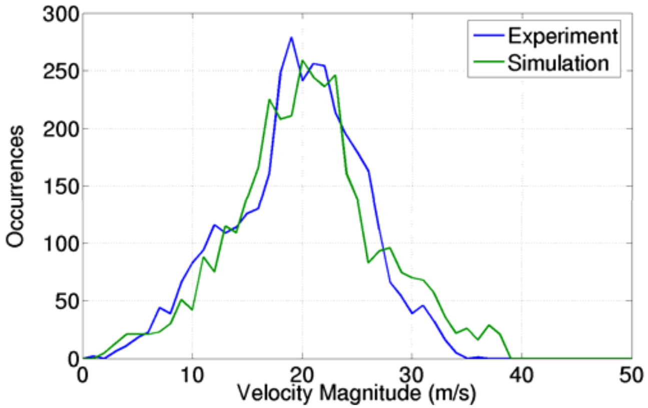


Gas Shock Tube Experimental Images (top)
Calculations using RAGE code (bottom)

- Gas shock tube experiments supply validation data for mix models
- Working towards quantitative metrics to compare the experimental measurements and the calculations
- Developing quantitative initial condition measurements



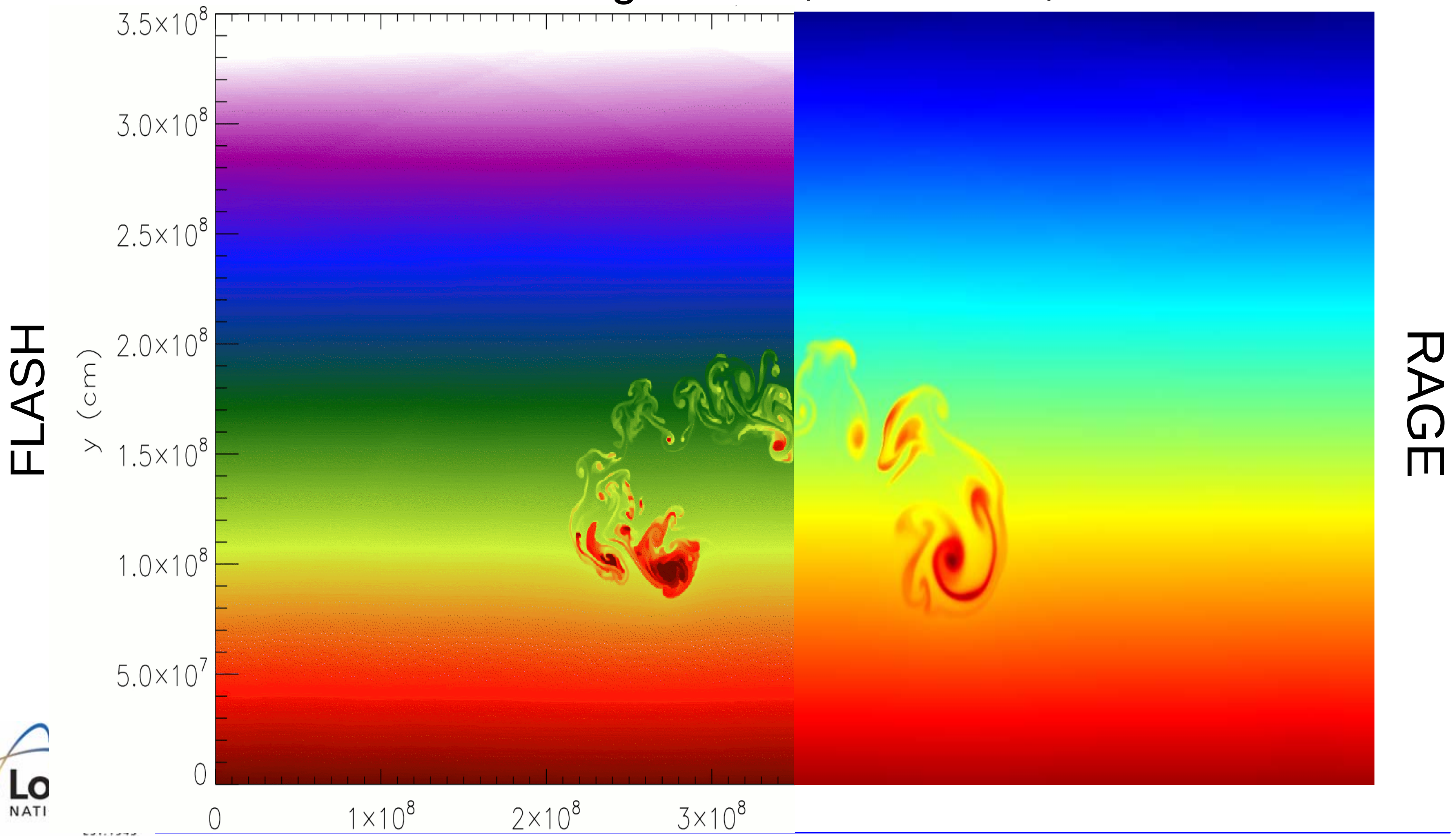
Vorticity values for experiment (left) and RAGE calculation (right)



Velocity Histograms for experiment (blue) vs. RAGE calculation (green)

"Rising Bubble" Code-Code Comparison in FLASH (U. Chicago) and RAGE (LANL)

Rising Bubble, 640x1280, 10s



Present V&V Challenges

- Hierarchical decomposition
  - Tightly coupled phenomena; Cost trade-offs
- Metrics for comparing simulation / experiment
  - Multi-dimensional data
  - Relevance to end-user applications
- Calibration vs. prediction and extrapolation
  - Underlying physical models must be credible

Drives Physics & Math Requirements

- Integrating Code Development/Testing
- Uncertainty quantification
  - Limits in computer resources
  - Large bounds from propagated uncertainties

*Drives ASC Computing **Capacity***

- Simulation Convergence / Error Estimates
  - Relevant test problems and convergence metrics
  - Statistical error-model development
  - Multi-Dimensional Error Estimate Requirements

*Drives ASC Computing **Capability***

ASC Computing Capacity Drivers

- **Born V&V'd** – Regular exercise and enlargement of V&V test suites
- **Uncertainty Analysis**

Computational simulations

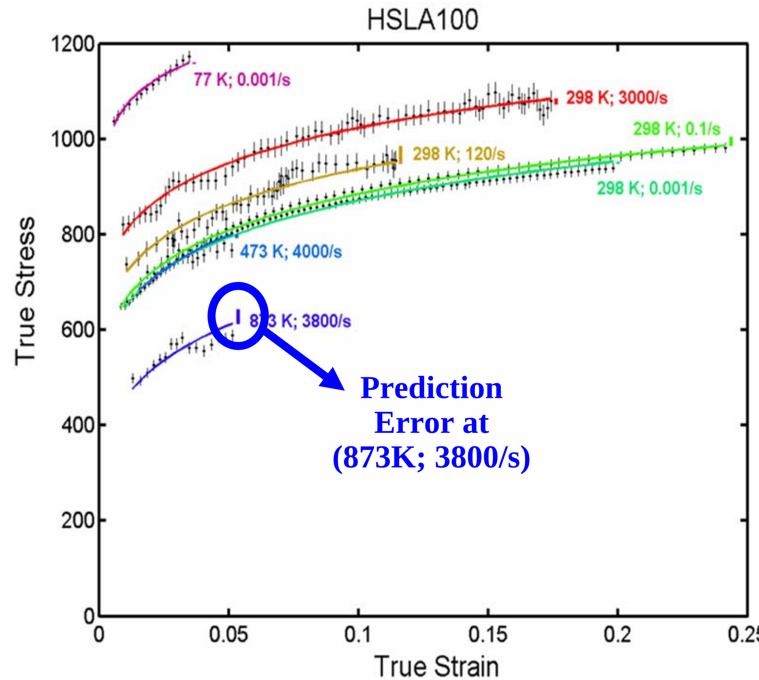
- Characterize/propagate input uncertainty
- Synthesis of experimental data

Experimental measurements

- Test-To-Test repeatability; Unit-to-Unit variability
- Computational processing of measurements

Test/analysis comparison

- Comparative metrics in the presence of numerical errors & numerical/physical uncertainty
- Calibration of uncertain model parameters
- Quantitative assessment of predictive accuracy under uncertainty



Comparisons of Hopkinson Bar Test Results with Zerilli-Armstrong Model for High-Strength Steel

V&V requires large-scale computing capacity

Tri-lab Verification Test Suite Coverage in a Crestone Project Code

Problem Name	1D Uniform	1D AMR	1D timestep	1D auto	2D uniform	2D AMR	2D auto	3D uniform	3D AMR	3D auto
Su-Olson	green	yellow	yellow	yellow	yellow	yellow	yellow	yellow	yellow	yellow
Cog-8	green	yellow	blue	yellow	blue	blue	blue	blue	blue	blue
Mader	green	yellow	yellow	yellow	yellow	yellow	yellow	yellow	yellow	yellow
RMTV	yellow	yellow	blue	yellow	blue	blue	blue	blue	blue	blue
Noh	green	yellow	blue	yellow	green	blue	blue	blue	blue	blue
Sedov	green	yellow	blue	yellow	green	blue	blue	blue	blue	blue
Sood	green	yellow	yellow	yellow	blue	blue	blue			

- green: FY04 (Kamm, Kirkpatrick)
- yellow: FY05 (Timmes, Gisler, Hrbek)
- blue: FY06 (Timmes, Fryxell, Hrbek)

Caveats:
- Tri-lab suite is a subset of desired full test suite
- Should measure apps-specific code coverage

Question: How many ASC codes have done this?

Born V&V'd is driving code-test requirements, which drives computing capacity requirements

ASC Computing Capability Drivers: Code/Calculation Verification as Example

Simulation Convergence / Error Estimates

- Quantified w/ Code & Calculation Verification

Code Verification

Test Problems
w/ Analytical Solns

- The process of gathering evidence that the computer implementation of the mathematical model and its associated solution are correct.
- "Solving the equations correctly"

Calculation Verification

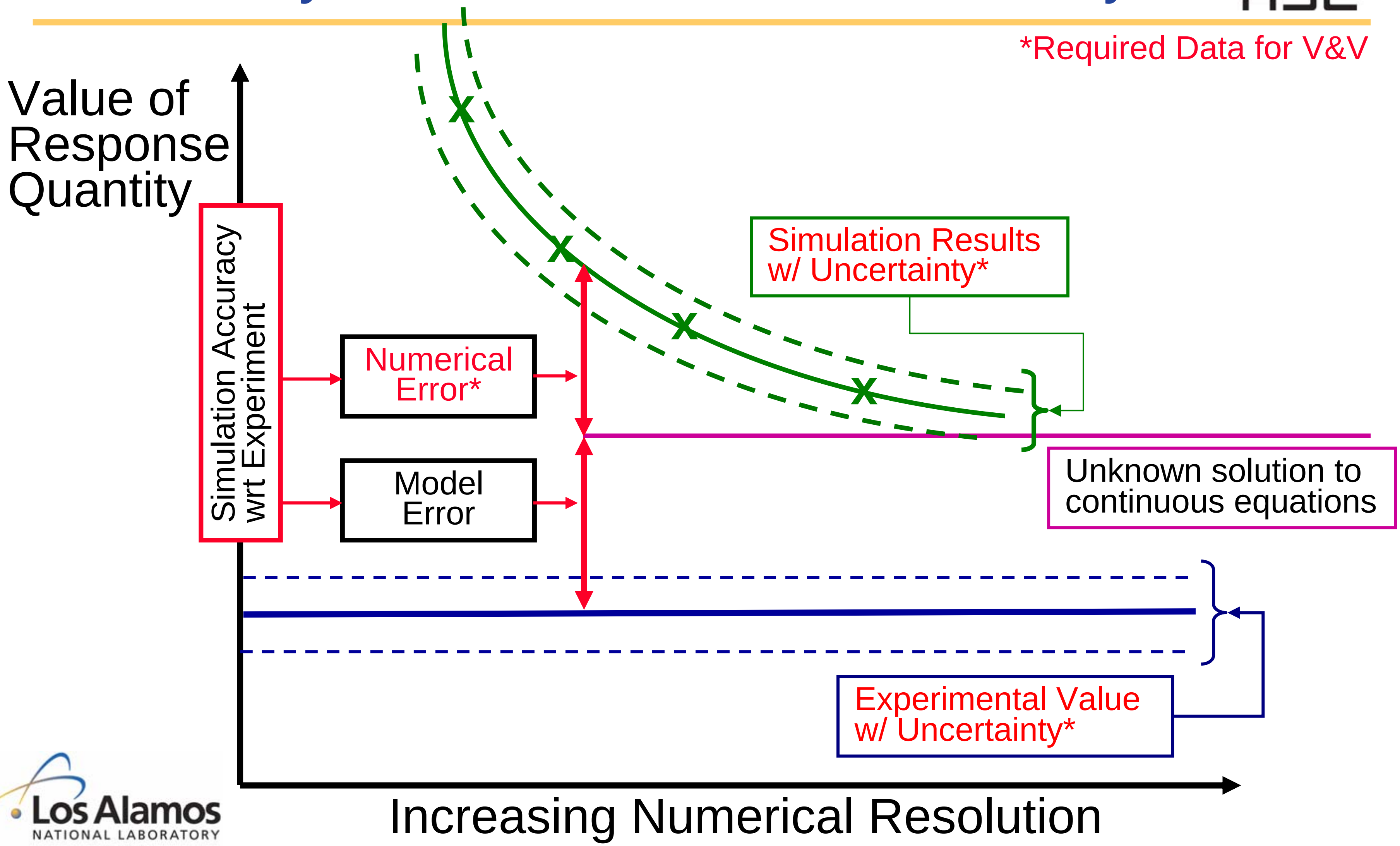
Test Problems
w/o Analytical Solns

- The process of assessing convergence & estimating numerical error for 'real' problems.
- "Assessing numerical error to supplement numerical uncertainty"

V&V requires large-scale computing capability

Notional Definitions of Uncertainty, Numerical Error and Accuracy

*Required Data for V&V

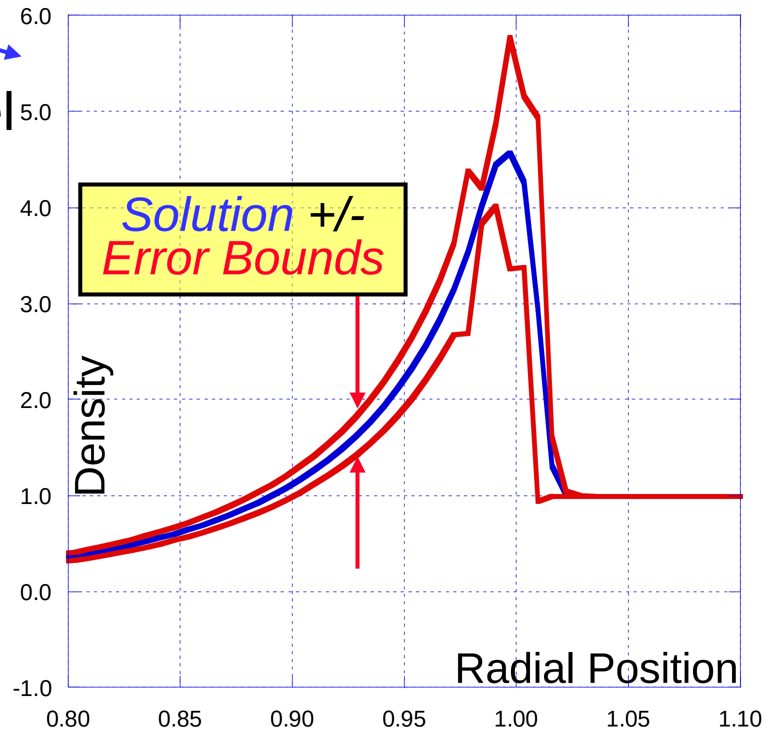


Calculation Verification: Estimating Numerical Errors for Relevant Problems

- **Objective**: Estimate numerical errors for multi-dimensional relevant simulations
  - scalar, integrated simulation results
  - convergence assessment and error estimation for simulation fields
- **Process**: Sensitivity studies (error model development) and grid-convergence studies (numerical error estimate)
  - *requires few, massive simulations*
- **Example**: Sedov Blast-Wave Problem

V&V requires large-scale computing capability

Snapshot of the Sedov Blast Wave Problem Simulation Results w/ Numerical Error Estimates



LANL ASC V&V Program: Conclusion

- **V&V Motivation**: Quantitative measures of
  - **Scientific credibility**: understanding the degree to which the code produces the **right answer for the right reason**
  - **Scientific utility**: ensuring that the code can address the **right questions** regarding complex scientific phenomena
  - Getting the right answer to the right questions for the right reason is a prerequisite for **predictive capability**

- **V&V Requirements**:
  - Hierarchical decomposition / Test Problems
  - Analysis Methods / Relevant simulation metrics
  - Integrated code development / testing system
  - Sufficient computing capacity/capability

V&V: Quantitative Bridge Between Physics Models, Simulation & Experiment

Abstract

This presentation provides an introduction and overview of the LANL ASC Verification and Validation program, including an explanation of the program mission, structure, the role in the software life cycle, and a description of hierarchical verification and validation assessments with an example.