Unclassified, Unlimited Release
SAND2007-2483C

Verification & Validation:
Measured Credibility, on Demand, for Stockpile Applications

Martin Pilch PhD
SNL/ASC/V&V Program Element Manager
QMU & Management Support, Dept. 1221
mpilch@sandia.gov

Presented at:
High Speed Computing Conference
Salishan Lodge in Gleneden Beach, Oregon
April 23-26, 2007

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company,
for the United States Department of Energy under contract DE-AC04-94AL85000.

We Need V&V Because We Care!

- **Stockpile applications**
  - **There *is* a decision context: High consequence design and decision making associated with nuclear weapons**
- **On Demand**
  - **Agility and responsiveness are critical to the nuclear weapons complex of the future**
- **Measured credibility**
  - **What's the “certificate of credibility” that accompanies M&S results**

M&S Increasingly Contributes to *Risk-Informed Decisions* at Sandia

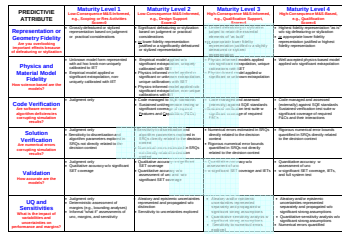
Stockpile Lifecycle Risk-Informed Decisions:
Assertion-Based and Challenged via Peer Review

Testing

Play Movie

CF=M/U

PCMM



PREDICTIVE ATTRIBUTE	Maturity Level 1 (e.g., Research or Pre-Decisional Support)	Maturity Level 2 (e.g., Design Support Quality)	Maturity Level 3 (e.g., Qualification Support, Required)	Maturity Level 4 (e.g., Qualification Support, Desired)
Representation or Geometry Fidelity (Are you overlooking important features or oversimplifying or idealizing or replicating?)	Grossly simplified or ideal representation of geometry or practical considerations.	Significant debeaming or replication of features or practical considerations. ▶ lower fidelity representation ▶ Lower fidelity representation or ideal representation	Detailed representation of geometry, including all features and elements of as built geometrical lower fidelity representation or ideal representation or replicated	▶ Highest fidelity representation: as-is or as-built geometry, including ▶ appropriate lower fidelity representation (justified as highest fidelity representation)
Physics and Material Model Fidelity (How science-based are the models?)	Ultra-low model from representative data, or from non-varying calibrated SET. Empirical model applied in significant extrapolation with uncertainty unknown with SET	Empirical model applied in significant extrapolation with uncertainty calibrated with SET	Physics-based models applied with engineering judgment, unique calibration with SET. Physics-informed model applied in significant extrapolation with SET	Well accepted physics-based model applied with significant extrapolation
Code Verification (Are software errors or algorithm deficiencies corrupting the results?)	Judgment only	Code managed in SCM, standard regression tests with significant coverage of required features and capabilities (FRCs)	Code managed in SCM, standard regression tests with 100% coverage of required FRCs	Code managed and automated regression tests with 100% coverage of required FRCs and their interactions
Solution Verification (Are numerical errors corrupting the results?)	Judgment only. Qualitative accuracy w/o significant SET coverage	Numerical error estimation (e.g. discretization error) or bounding (e.g. iterative error) in SRQs not directly related to the decision context	Numerical error estimation (e.g. discretization error) or bounding (e.g. iterative error) in SRQs directly related to the decision context	Numerical errors estimated in SRQs directly related to the decision context. Regression numerical error bounds quantified in SRQs not directly related to the decision context
Validation (How accurate are the models?)	Judgment only. Qualitative accuracy w/o significant SET coverage	Quantitative accuracy w/o significant SET coverage. Quantitative accuracy w/o significant SET coverage	Quantitative accuracy with significant SET coverage and IETs	Quantitative accuracy, at experimental scale, with significant SET coverage, IETs, and full system test
UQ and Sensitivities (What is the impact of uncertainties on system performance and margins?)	Judgment only. Qualitative assessment of margins (e.g., operating analyzed in tactical "what if" assessments of sens, margins, and sensitivity	Aleatory and epistemic uncertainties characterized and propagated via dimension. Sensitivity to uncertainties explored	Aleatory and epistemic uncertainties characterized and propagated via separately and segregated regression strong assumptions. Sensitivity to uncertainties explored with strong assumptions. Sensitivity to numerical errors	Aleatory and/or epistemic uncertainties characterized separately and propagated via significant strong assumptions. Sensitivity to uncertainties with significant strong assumptions. Numerical errors quantified

DECISION SUPPORT

C³:
- Conservative Requirements
- Conservative Scenarios
- Conservative Assessments

Decision Paradigm Change

QMU:
Quantified Margins and Uncertainties BE+U

Test/Experiment
Relevant Physical Data

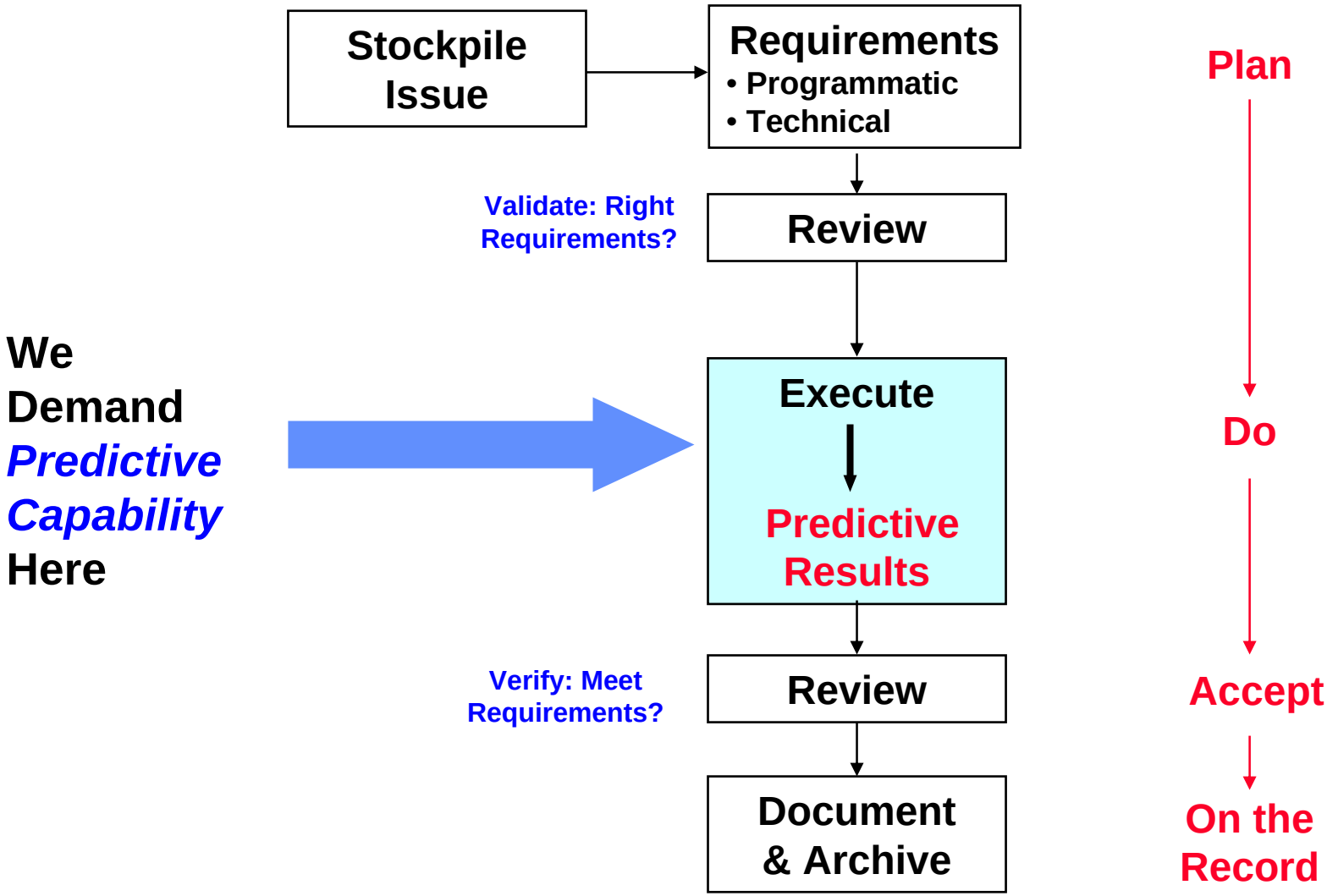
M&S
Predictive Capability

CREDIBILITY
Of Models and Experiments

For improved technical basis, communication, decisions

Concepts of Stockpile Computing

SAND2004-2479



What Does it Mean "to Predict"?

American Heritage Dictionary:

- **Predict: To state, tell about, or make known in advance, especially on the basis of *special knowledge****

What *special knowledge* do we demand of M&S to assert a predictive capability?

***A CS&E prediction is a M&S-based evaluation prior to or in lieu of physical measurement**

You Can't Measure and Communicate "it" Unless You Know What "it" Is

- **Some Attributes of Predictive Capability**
  - **Representation or geometric fidelity**
  - **Physics and material model fidelity (predictive science)**
  - **Code verification**
  - **Solution verification**
  - **Validated models**
  - **Uncertainty quantification with sensitivity analysis**

How Much is Enough?

- Sufficiency can only be discussed in an application context
- Graded approach
  - Maturity Level 0:
    - Low-consequence, M&S-informed (e.g., scoping studies)
  - Maturity Level 1:
    - Low-consequence, M&S-informed (e.g., design support)
  - Maturity Level 2:
    - High-consequence, M&S-informed (e.g., qual support)
  - Maturity Level 3:
    - High-consequence, M&S-based (e.g., qualification)
- Negotiate expectations for future work or communicate maturity for work already done

Increasing Rigor Expected

This is Where We Are Going

Predictive Capability Maturity Model (PCMM)

MATURITY ATTRIBUTE	Maturity Level 0 Low Consequence, Minimal M&S Impact, e.g. Scoping Studies	Maturity Level 1 Moderate Consequence, Some M&S Impact, e.g. Design Support	Maturity Level 2 High-Consequence, High M&S Impact, e.g. Qualification Support	Maturity Level 3 High-Consequence, Decision-Making Based on M&S, e.g. Qualification or Certification
Representation and Geometric Fidelity Are representation errors corrupting simulation conclusions?	• Judgment only • Little or no representational or geometric fidelity for the system and BCs	• Significant simplification or stylization of the system and BCs • Geometry or representation of major components is defined	• Limited simplification or stylization of major components and BCs • Geometry or representation is well defined for major components and some minor components	• Essentially no simplification or stylization of components in the system and BCs • Geometry or representation of all components is at the detail of "as built", e.g., gaps, material interfaces, fasteners, welds, adhesive bonding, surface finish
Physics and Material Model Fidelity How fundamental are the physics and material models and what is the level of model calibration?	• Unknown model form • Unknown, unjustified, or significant extrapolation • No coupling of models	• Empirical model form • Moderate extrapolation • Minimal or ad hoc coupling of models	• Physics-informed models calibrated to relevant SET/IET data • Minimal extrapolation • One-way coupling of models	• Physics-based models minimal need for calibration using SETs and IETs • Interpolation • Full, two-way, coupling of models
Code Verification Are software errors or algorithm deficiencies corrupting the simulation results?	• Judgment only • Minimal testing of any software elements • Little or no SQE procedures specified or followed	• Codes managed by SQE procedures • Unit and regression testing routinely conducted with significant code coverage	• Codes managed by SQE procedures and assessed by organization • Sustained Verification Test Suites (VERTS) regularly w significant functionality coverage	• SQE procedures reviewed by independent, external panel • Sustained Verification Test Suites (VERTS) regularly w significant coverage of functionality and their interactions
Solution Verification Are human procedural errors or numerical solution errors corrupting simulation conclusions?	• Judgment only, numerical errors have an unknown effect on simulation results • Input/output not verified	• Sensitivity to discretization and algorithm parameters explored for some System Response Quantities (SRQs) • Input/output verified only by the analysts	• Numerical errors estimated for discretization and algorithm parameters for relevant SRQs • Input/output data independently verified internal to the organization	• Numerical errors rigorously quantified for all relevant SRQs • Input/output data independently verified externally to the organization
Model Validation How accurate are the simulation results at various tiers in a validation hierarchy?	• Judgment only • Few, if any, comparisons with relevant SETs or IETs	• Qualitative assessment of accuracy • Model comparisons to SETs w sig coverage w sig coverage of physics	• Quantitative assessment of model accuracy w/o model or measurement uncertainty • Model comparisons to SETs and IETs w sig coverage of physics and limited coverage of physics interactions	• Quantitative assessment of model accuracy w model and measurement uncertainty • Model comparisons to SETs, IETs, and full system tests (as appropriate) w sig coverage of physics and their interactions
Uncertainty Quantification and Sensitivity Analysis What is the impact of variabilities and uncertainties on system performance and margins?	• Judgment only or unc not addressed • Sensitivities to unc not addressed • Numerical (propagation) errors not addressed	• Aleatory and epistemic (A&E) unc not segregated on inputs/outputs w/o proper interpretation • Qualitative sensitivity to some unc explored w informal "what if" studies • Sensitivity to numerical (propagation) errors estimated for UQ results • Many strong assumptions	• A&E unc segregated on inputs/outputs w proper interpreted • Quantitative sensitivity analyses conducted for some uncertainties • Quantitative sensitivity to unc estimated • Sensitivity to numerical (propagation) errors estimated for UQ/SA • Some strong assumptions	• A&E uncertainties due to all plausible environments and scenarios of the system are analyzed • Quantitative sensitivity to unc rigorously bounded • Numerical (propagation) errors rigorously quantified for UQ/SA • No significant strong assumptions

Why PCMM?

- **Educate the decision maker about what should be expected from M&S**
- **Measure/communicate process maturity (not adequacy of results) associated with M&S in a decision context**
- **Provide program vision so that technical and infrastructure needs can be leveraged across multiple funding lines to enhance the credibility of M&S results**
- **Speak to the *whats*, not dictate the *hows***

Setting the National Agenda in V&V

Review of the Literature

- **Capability Maturity Model Integration (CMMI) developed by the Software Engineering Institute, Carnegie Mellon University:**
  - **Developed maturity assessment procedures for a wide range of development processes, but not specifically for M&S**
- **DoD/NASA Technology Readiness Levels (TRLs):**
  - **Appropriate for assessing the risk of introducing new technologies into new engineering systems or existing hardware**
- **Harmon and Youngblood (DoD/DMSO) developed a process assessing the maturity of five contributing elements for M&S:**
  - **Conceptual model of the simulation**
  - **Verification results from the contributing products**
  - **Appropriateness of the referent**
  - **Appropriateness of the validation criteria**
  - **Adequacy of the simulation results for the intended use**
  - **Six levels of maturity**

Review of the Literature (continued)

- Logan and Nitta (LLNL) proposed a process to assess the maturity of verification and validation activities: Ver/Val Meters
- Pilch et al proposed a process to assess maturity of four elements of stockpile computing:
  - Qualified analysts
  - Qualified codes
  - Qualified computational infrastructure
  - Appropriate levels of formality
  - Four levels of maturity
- NASA recently adopted an engineering standard for M&S with seven contributing elements:
  - Code verification
  - Solution verification
  - Validation
  - Predictive capability
  - Technical review
  - Process control
  - Operator and analyst qualification
  - Four levels of maturity

Representation or Geometric Fidelity

Are representation limitations corrupting simulation results?

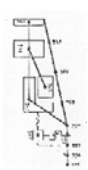
Little or no representational or geometric fidelity for system and BC's

Significant simplification and stylizations (S&S)

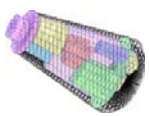
Limited simplification and stylizations (S&S)

"as built"

Increasing Rigor or Completeness →



10 years ago:
Shellshock 2D
NASTRAN
200 dof

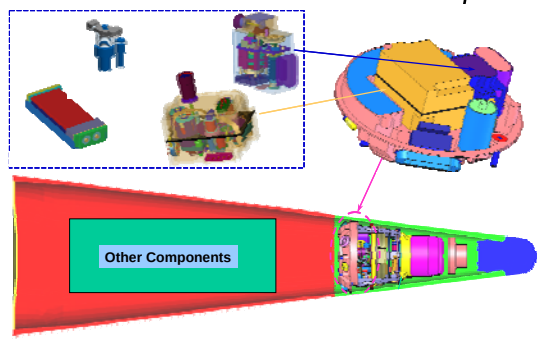


Recent Past:
NASTRAN

30,000 dof

800,000 dof

Today:
SALINAS MP
>10M dof

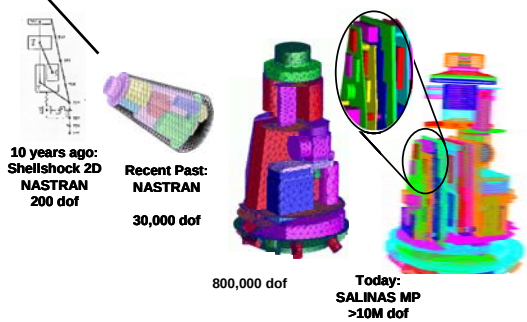


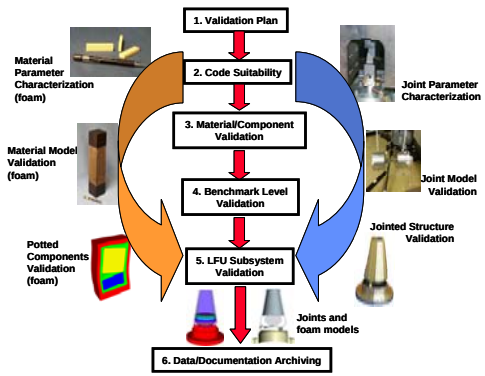
Physics and Material Model Fidelity

How fundamental are the physics and material models and what is the level of model calibration?

Unknown model form	Empirical model forms	Physics-informed models calibrated to relevant SET/IET	Physics-based models
Significant extrapolation	Moderate extrapolation	Minimal extrapolation	Interpolation
No coupling	Minimal or ad-hoc coupling	1-way coupling	2-way coupling

Increasing Rigor or Completeness →





Code Verification

Are software errors or algorithm deficiencies corrupting simulation results?

Judgment only

Code managed to SQE standards

Sustained unit/regression tests w sig coverage of F&C

SQE + assessment

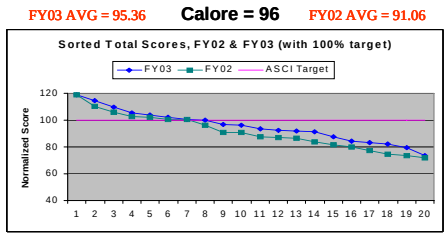
Sustained VERTS w sig function coverage

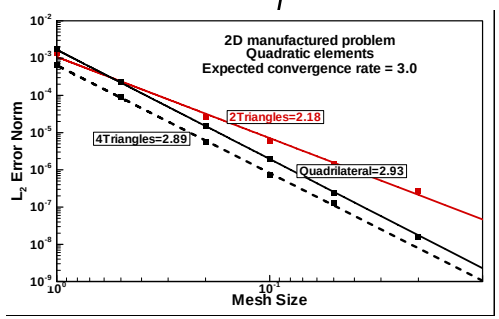
SQE(A)

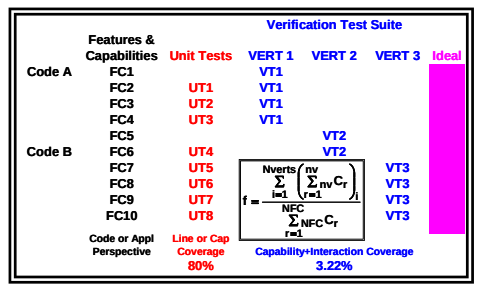
Sustained VERTS w sig coverage of function interactions

Increasing Rigor or Completeness →

Code/Code Comparisons







	Features & Capabilities	Unit Tests	VERT 1	VERT 2	VERT 3	Ideal
			Verification Test Suite			
Code A	FC1		VT1			
	FC2	UT1	VT1			
	FC3	UT2	VT1			
	FC4	UT3	VT1			
	FC5			VT2		
Code B	FC6	UT4		VT2		
	FC7	UT5			VT3	
	FC8	UT6			VT3	
	FC9	UT7			VT3	
	FC10	UT8			VT3	
	Code or Appl Perspective	Line or Cap Coverage 80%	Capability+Interaction Coverage 3.22%			

$$f = \frac{\sum_{i=1}^{Nverts}\left(\sum_{r=1}^{nv} C_r\right)_i}{\sum_{r=1}^{NFC} C_r}$$

Solution Verification

Are numerical or processing errors corrupting simulation results?

Judgment only

Sensitivity to discretization and algorithm parameters explored

Input/out verified only by analyst

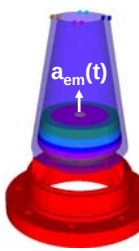
Estimate numerical errors

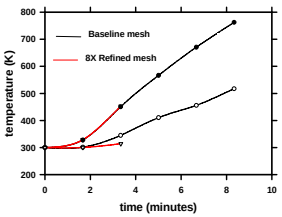
Independent verification of input/output internal to org

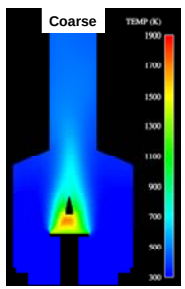
Quantify rigorous numerical error bounds

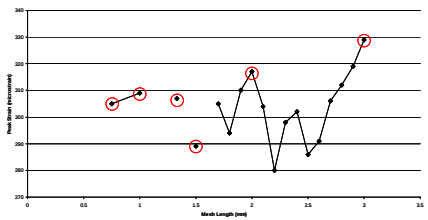
Independent verif of input/output ext to org

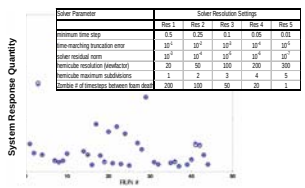
Increasing Rigor or Completeness →



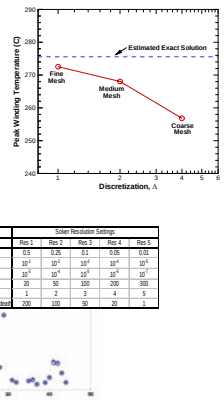


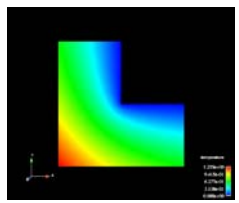






Solver Parameter	Solver Resolution Settings				
	Res 1	Res 2	Res 3	Res 4	Res 5
minimum time step	0.5	0.25	0.1	0.05	0.01
time-matching truncation error	10^{-5}	10^{-7}	10^{-9}	10^{-9}	10^{-9}
solver residual norm	10^{-5}	10^{-7}	10^{-9}	10^{-9}	10^{-9}
hemicube resolution (viewfactor)	20	50	100	200	300
hemicube maximum subdivisions	1	2	3	4	5
Tombe # of timesteps between form depth	200	100	50	20	1





Validation

How Accurate are the simulation results at various tiers of the validation hierarchy?

Judgment Only
Few comps to SETs or IETs

Qualitative accuracy assessment

Separate Effects Tests (SETs) w sig coverage of physics

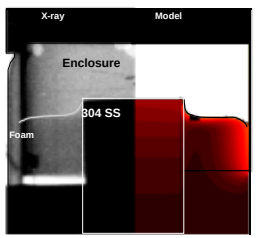
Quantitative accuracy assessment w/o model or meas unc

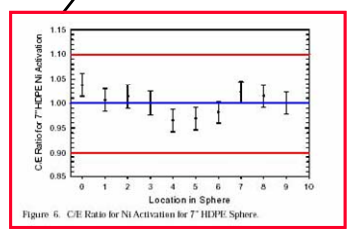
SET and Integral Effects Tests (IETs) w sig physics coverage and limited physics interactions

Quantitative accuracy assessment w model and meas unc

SET coverage, IETs, and full system tests w physics coverage and interaction coverage

Increasing Rigor or Completeness →



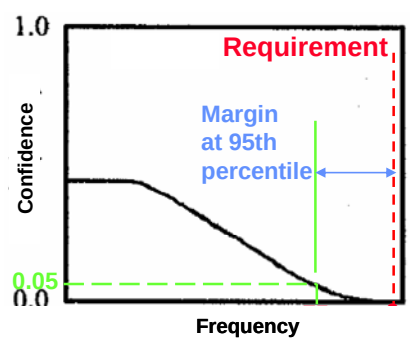


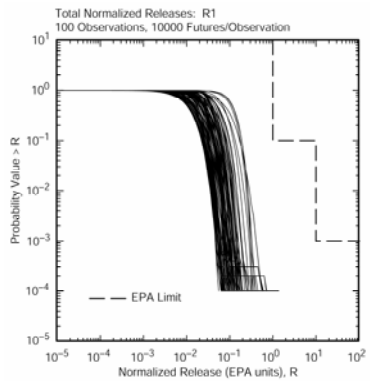
Uncertainty Quantification and Sensitivities

What is the impact of variabilities and uncertainties on performance and margins?

Judgment only Deterministic assessment of margins	Aleatory and epistemic (A&E) unc. interpreted on inputs/outputs w/o distinction Sensitivities to some unc explored Sensitivity to numerical errors explored	A&E unc segregated on inputs/outputs and properly interpreted w strong assumptions Sensitivities w strong assumptions Sensitivity to numerical errors estimated	A&E unc segregated on inputs/outputs and properly interpreted w/o strong assumptions Sensitivities w/o strong assumptions Numerical errors rigorously quantified

Increasing Rigor or Completeness →





ASC Projects Should Map to Capability/Agility Needs

MATURITY ATTRIBUTE	Maturity Level 0 Low Consequence, Minimal M&S Impact, e.g. Scoping Studies	Maturity Level 1 Moderate Consequence, Some M&S Impact, e.g. Design Support	Maturity Level 2 High-Consequence, High M&S Impact, e.g. Qualification Support	Maturity Level 3 High-Consequence, Decision-Making Based on M&S, e.g. Qualification or Certification
Geometric Fidelity Are representation errors corrupting simulation conclusions?	geometric fidelity for the system and BCs	• Geometry or representation of major components is defined	defined for major comp some minor componen	representation of all is at the detail of "as aps, material interfaces, elds, adhesive bonding, surface finish
Physics and Material Model Fidelity How fundamental are the physics and material models and what is the level of model calibration?	• Un • Un or ext • No coupling of models	cal model form ate extrapolation l or ad hoc coupling els	• Physics-informed models calibrated to relevant SET/IET data • Minimal extrapolation of models	• Physics-based models minimal need for calibration using SETs and IETs • Interpolation • Full, two-way, coupling of models
Code Verification Are software errors or algorithm deficiencies corrupting the simulation results?	• Judgment only • Minimal testing of software elements • Little or no SQE procedures specified or followed	• Codes managed by SQE s egression testing onducted with significant code coverage	by SQE edures and assessed by ization ained Verification Test Suites TS) regularly w significant onality coverage	• SQE procedures reviewed by independent, external panel • Sustained Verification Test Suites (VERTS) regularly w significant coverage of functionality and their interactions
Solution Verification Are human procedural errors or numerical solution errors corrupting simulation conclusions?	• Judgment only, numerical errors have an unknown effect on simulation results • Input/output not verified	• Sensitivity to discretization and algorithm parameters explored for some System Response Quantities (SRQs) • Input/output verified only by the analysts	erical errors estimated for etization and algorithm neters for rele /output data in ed internal to t	• Numerical errors rigorously quantified for all relevant SRQs dependently the organization
Model Validation How accurate are the simulation results at various tiers in a validation hierarchy?	• Judgment only • Few, if any, comparisons with relevant SETs or IETs	• Qualitative assessment of accuracy • Model comparis w sig coverage coverage of phy	• Quantitative assessment of model y ETs and hysics and limited coverage of physics interactions	• Quantitative assessment of model accuracy w model and measurement uncertainty • Model comparisons to SETs, IETs, and full system tests (as appropriate) w sig coverage of physics and their interactions
Uncertainty Quantification and Sensitivity Analysis What is the impact of variabilities and uncertainties on system performance and margins?	• Judgment only or unc not addressed • Sensitivities to unc not addressed • Numerical (propagation) errors not addressed	• Aleatory and epistemic (A&E) unc not segregated on in interp • Quali some informal "what if" studies • Sensitivity to numerical (propagation) errors estimated for UQ results • Many strong assumptions	• A&E unc segregated on inputs/outputs w proper interpreted Quantit conduc Quantit estimat • Sensitivity to numerical (propagation) errors estimated for UQ/SA • Some strong assumptions	• A&E uncertainties due to all plausible environments and scenarios of the system are analyzed • Quantitative sensitivity to unc rigorously bounded • Numerical (propagation) errors rigorously quantified for UQ/SA • No significant strong assumptions

DART

Wisdom

PEM

SQE

Coda

Adaptivity

ValMetrics

Dakota

SAT

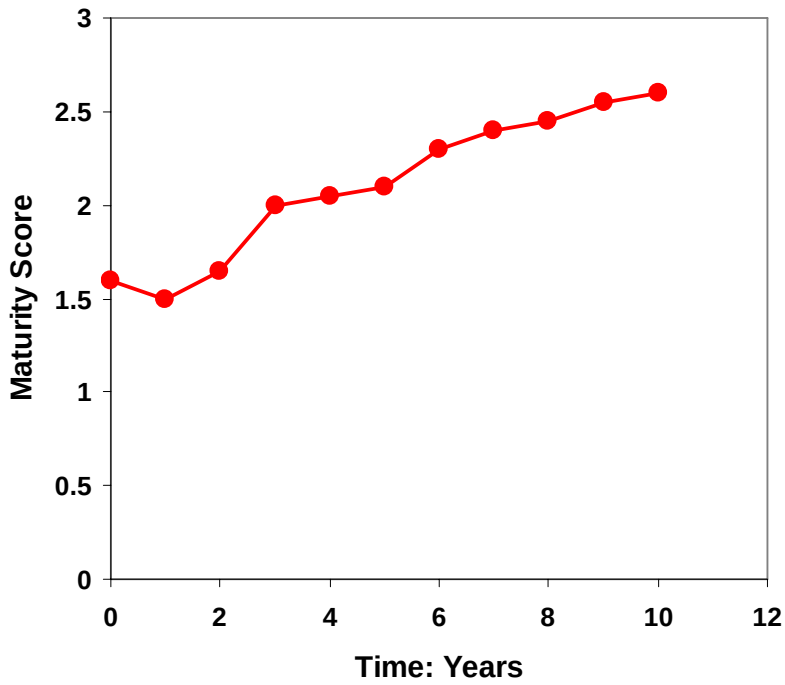
Predictive Capability Maturity Model (PCMM)

Communicating Credibility

MATURITY ATTRIBUTE	Maturity Level 0 Low Consequence, Minimal M&S Impact, e.g. Scoping Studies	Maturity Level 1 Moderate Consequence, Some M&S Impact, e.g. Design Support	Maturity Level 2 High-Consequence, High M&S Impact, e.g. Qualification Support	Maturity Level 3 High-Consequence, Decision-Making Based on M&S, e.g. Qualification or Certification
Representation and Geometric Fidelity Are representation errors corrupting simulation conclusions?	• Judgment only • Little or no representational or geometric fidelity for the system and BCs	• Significant simplification or stylization of the system and BCs • Geometry or representation of major components is defined	• Limited simplification or stylization of major components and BCs • Geometry or representation is well defined for major components and some minor components	• Essentially no simplification or stylization of components in the system and BCs • Geometry or representation of all components is at the detail of "as built", e.g., gaps, material interfaces, fasteners, welds, adhesive bonding, surface finish
Physics and Material Model Fidelity How fundamental are the physics and material models and what is the level of model calibration?	• Unknown model form • Unknown, unjustified, or significant extrapolation • No coupling of models	• Empirical model form • Moderate extrapolation • Minimal or ad hoc coupling of models	• Physics-informed models calibrated to relevant SET/IET data • Minimal extrapolation • One-way coupling of models	• Physics-based models minimal need for calibration using SETs and IETs • Interpolation • Full, two-way, coupling of models
Code Verification Are software errors or algorithm deficiencies corrupting the simulation results?	• Judgment only • Minimal testing of any software elements • Little or no SQE procedures specified or followed	• Codes managed by SQE procedures • Unit and regression testing routinely conducted with significant code coverage	• Codes managed by SQE procedures and assessed by organization • Sustained Verification Test Suites (VERTS) regularly w significant functionality coverage	• SQE procedures reviewed by independent, external panel • Sustained Verification Test Suites (VERTS) regularly w significant coverage of functionality and their interactions
Solution Verification Are human procedural errors or numerical solution errors corrupting simulation conclusions?	• Judgment only, numerical errors have an unknown effect on simulation results • Input/output not verified	• Sensitivity to discretization and algorithm parameters explored for some System Response Quantities (SRQs) • Input/output verified only by the analysts	• Numerical errors estimated for discretization and algorithm parameters for relevant SRQs • Input/output data independently verified internal to the organization	• Numerical errors rigorously quantified for all relevant SRQs • Input/output data independently verified externally to the organization
Model Validation How accurate are the simulation results at various tiers in a validation hierarchy?	• Judgment only • Few, if any, comparisons with relevant SETs or IETs	• Qualitative assessment of accuracy • Model comparisons to SETs w sig coverage w sig coverage of physics	• Quantitative assessment of model accuracy w/o model or measurement uncertainty • Model comparisons to SETs and IETs w sig coverage of physics and limited coverage of physics interactions	• Quantitative assessment of model accuracy w model and measurement uncertainty • Model comparisons to SETs, IETs, and full system tests (as appropriate) w sig coverage of physics and their interactions
Uncertainty Quantification and Sensitivity Analysis What is the impact of variabilities and uncertainties on system performance and margins?	• Judgment only or unc not addressed • Sensitivities to unc not addressed • Numerical (propagation) errors not addressed	• Aleatory and epistemic (A&E) unc not segregated on inputs/outputs w/o proper interpretation • Qualitative sensitivity to some unc explored w informal "what if" studies • Sensitivity to numerical (propagation) errors estimated for UQ results • Many strong assumptions	• A&E unc segregated on inputs/outputs w proper interpreted • Quantitative sensitivity analyses conducted for some uncertainties • Quantitative sensitivity to unc estimated • Sensitivity to numerical (propagation) errors estimated for UQ/SA • Some strong assumptions	• A&E uncertainties due to all plausible environments and scenarios of the system are analyzed • Quantitative sensitivity to unc rigorously bounded • Numerical (propagation) errors rigorously quantified for UQ/SA • No significant strong assumptions

PCMM Can Be Used To Measure Progress In Predictive Capability Over Time

Application: Legacy Weapon in Fuel Fire	Avg/Avg Score	Avg/Min Score
Rollup Scores	1.6	1.1
Representation or Geometric Fidelity	2.8	2.0
Fire Environment	2.0	
Weapon	3.5	
Physics and Material Modelel Fidelity	1.6	0.5
Fire Environment	1.6	1.0
Fuel vaporization from spill	1.0	
Fluid mechanics	2.0	
Turbulent mixing	1.5	
Combustion	2.0	
Emmission	1.0	
Radiative transport to weapon	2.0	
Convective transport to weapon	2.0	
Weapon Thermal Response	1.7	0.0
Code Verification	1.5	1.5
Fire Environment	1.5	
Weapon Thermal Response	1.5	
Solution Verification	0.8	0.5
Fire environment	0.5	
Weapon thermal response	1.0	
Validation	1.1	0.0
Fire Environment	0.9	0.0
Weapon Thermal Response	1.2	0.0
UQ/SA	2.0	2.0



Measured Credibility, on Demand, for Stockpile Applications

- Decision makers need to understand predictive capability in order to make informed decisions and to efficiently leverage and make use of research dollars
- Progress in predictive capability needs to be measured in each individual decision context
  - Predictive capability is more than geometric fidelity or even physics fidelity
  - There is a need to define sufficiency (or adequacy) in each attribute of predicative capability
- The Predictive Capability Maturity Model provides a graded approach to assessing and measuring predictive capability for specific applications

The Credibility of M&S is Critical

"Due diligence means asking *all* the questions, even if you don't think you'll like the answers."

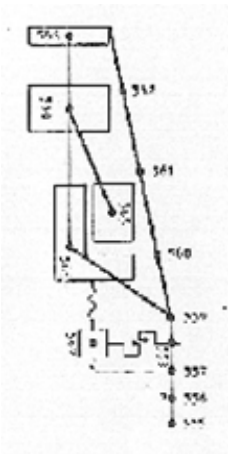
Representational (Geometric) Fidelity

Hyperlinks

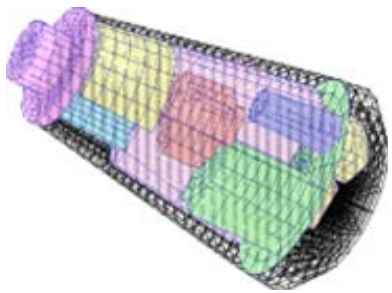
Imagine the Future!

Computing Speed - Dec. 2006

Progress in Representational Fidelity in Structural Dynamics

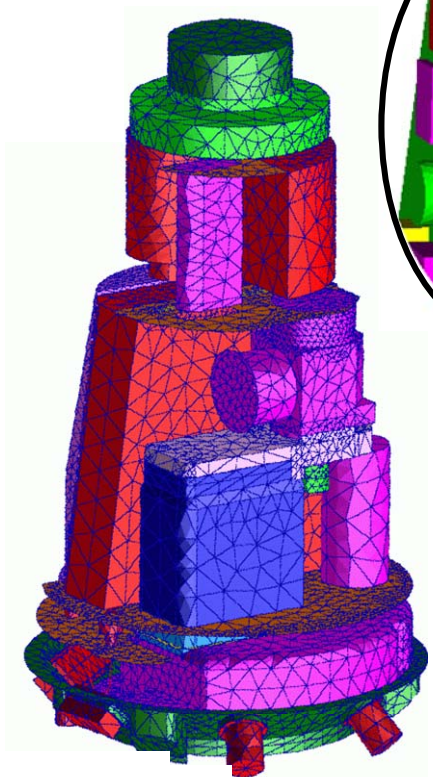


10 years ago:
Shellshock 2D
NASTRAN
200 dof

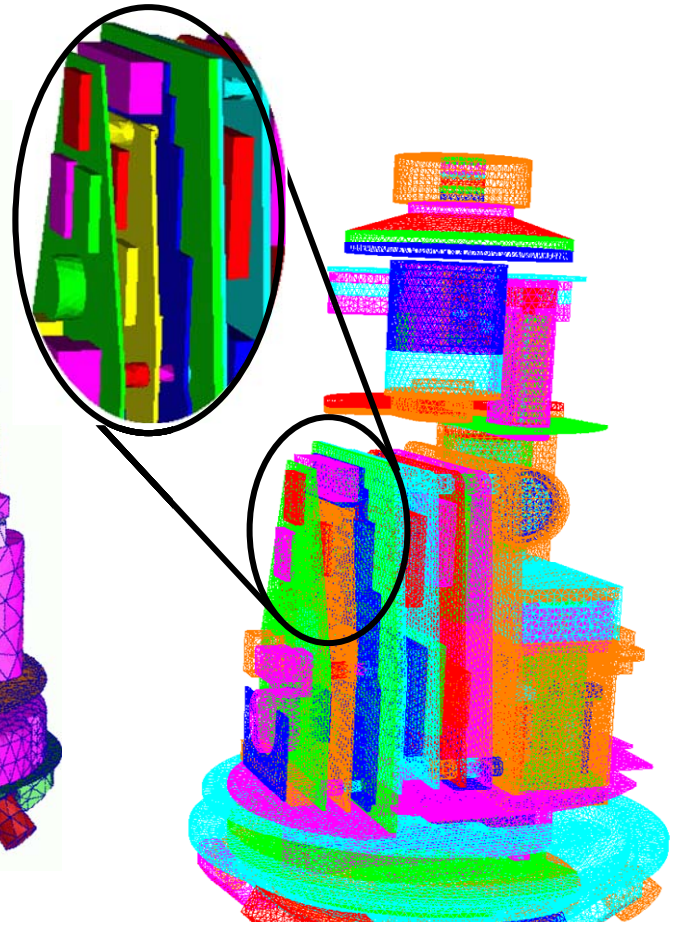


Recent Past:
NASTRAN

30,000 dof

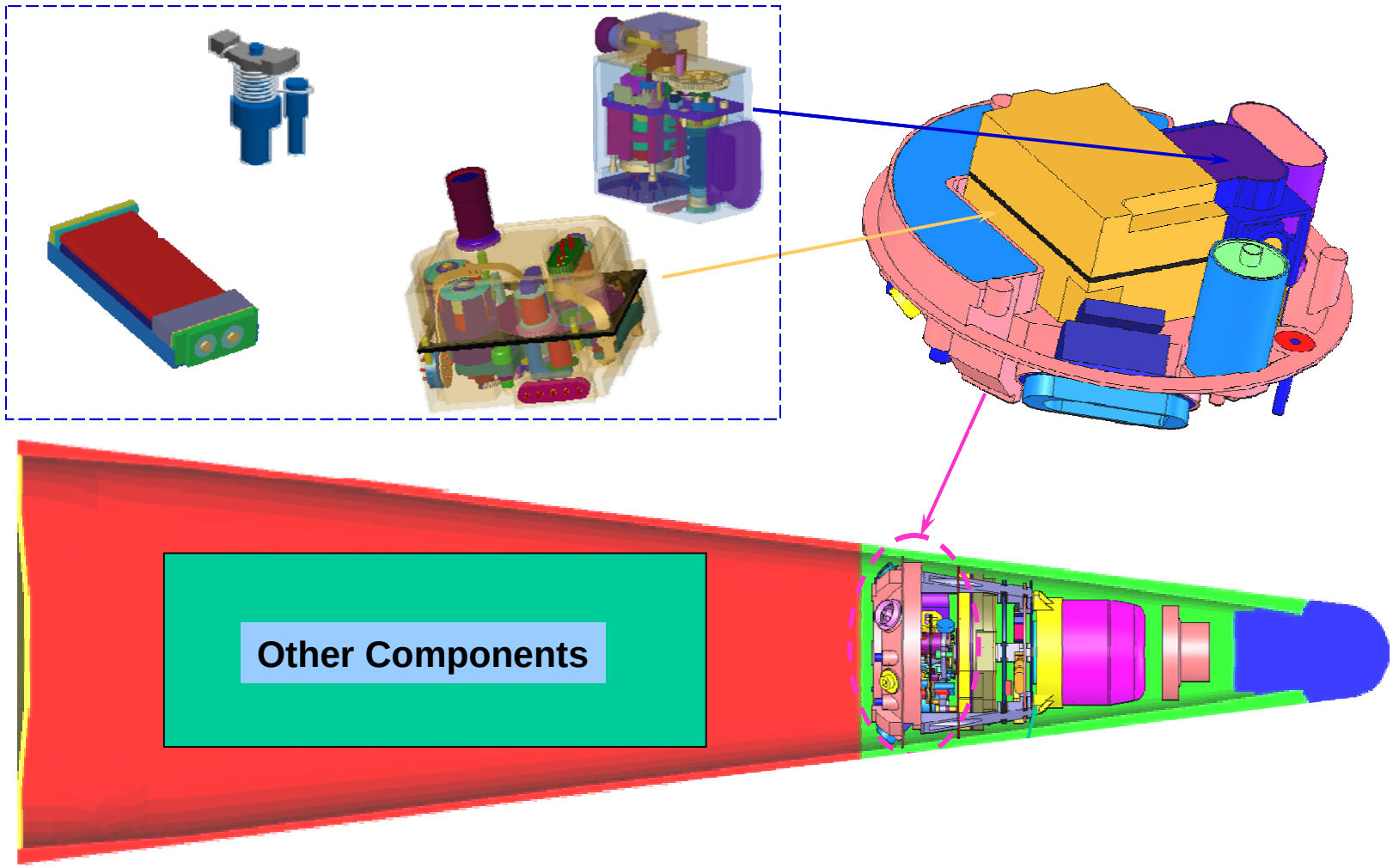


800,000 dof



Today:
SALINAS MP
>10M dof

Progress in Representational Fidelity Thermal Modeling



Physics Fidelity

Hyperlinks

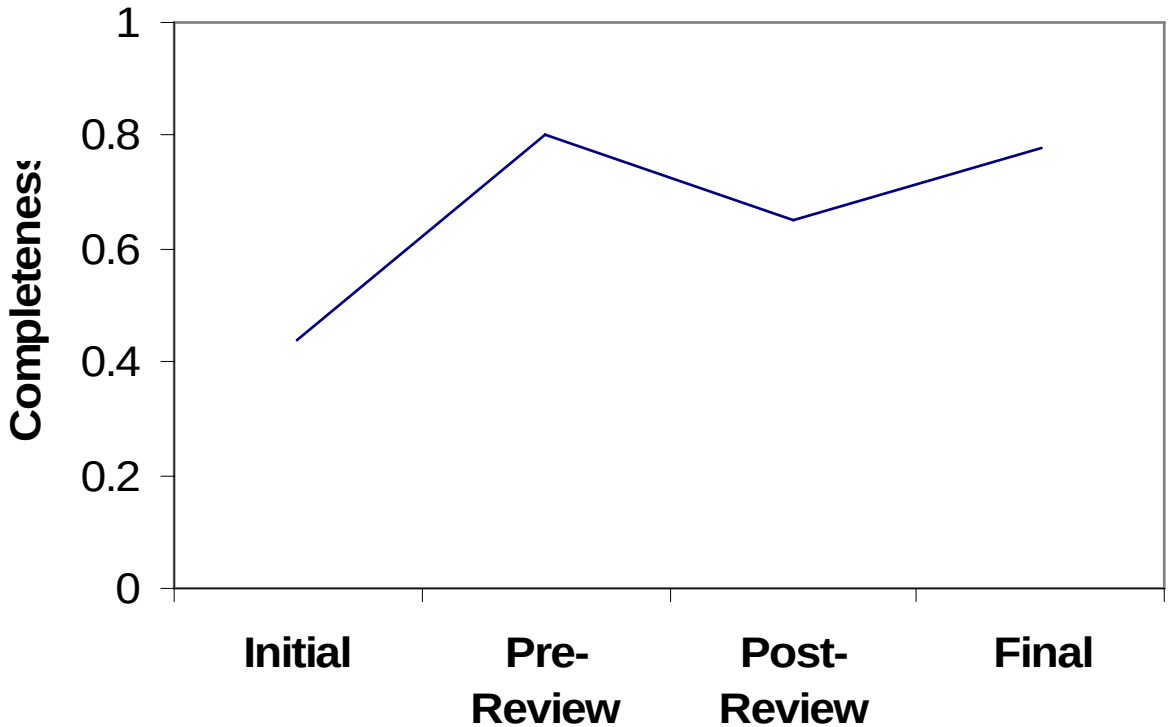
Phenomena Identification and Ranking Tables (PIRT)

Establish efficiency and sufficiency of activities

Phenomena	Importance	Adequacy: Model	Adequacy: Code	Adequacy: Validation
P1	H	H	M	L
P2	M	M	L	L
P3	L	L	L	L

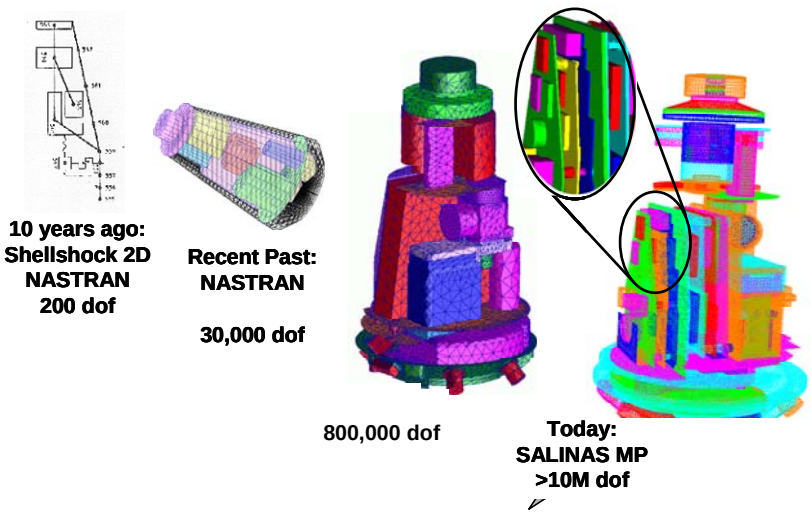
Gap = 5

Completeness = 0.44

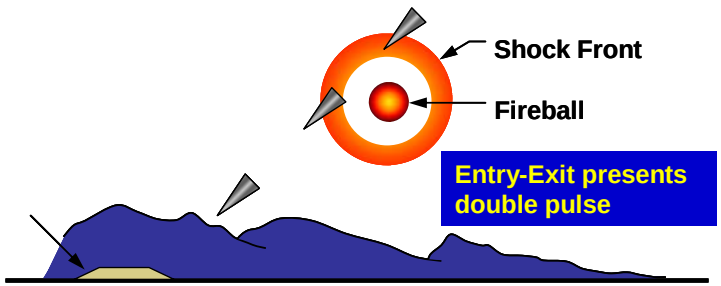


Low Physics Fidelity

- Conduct blast test

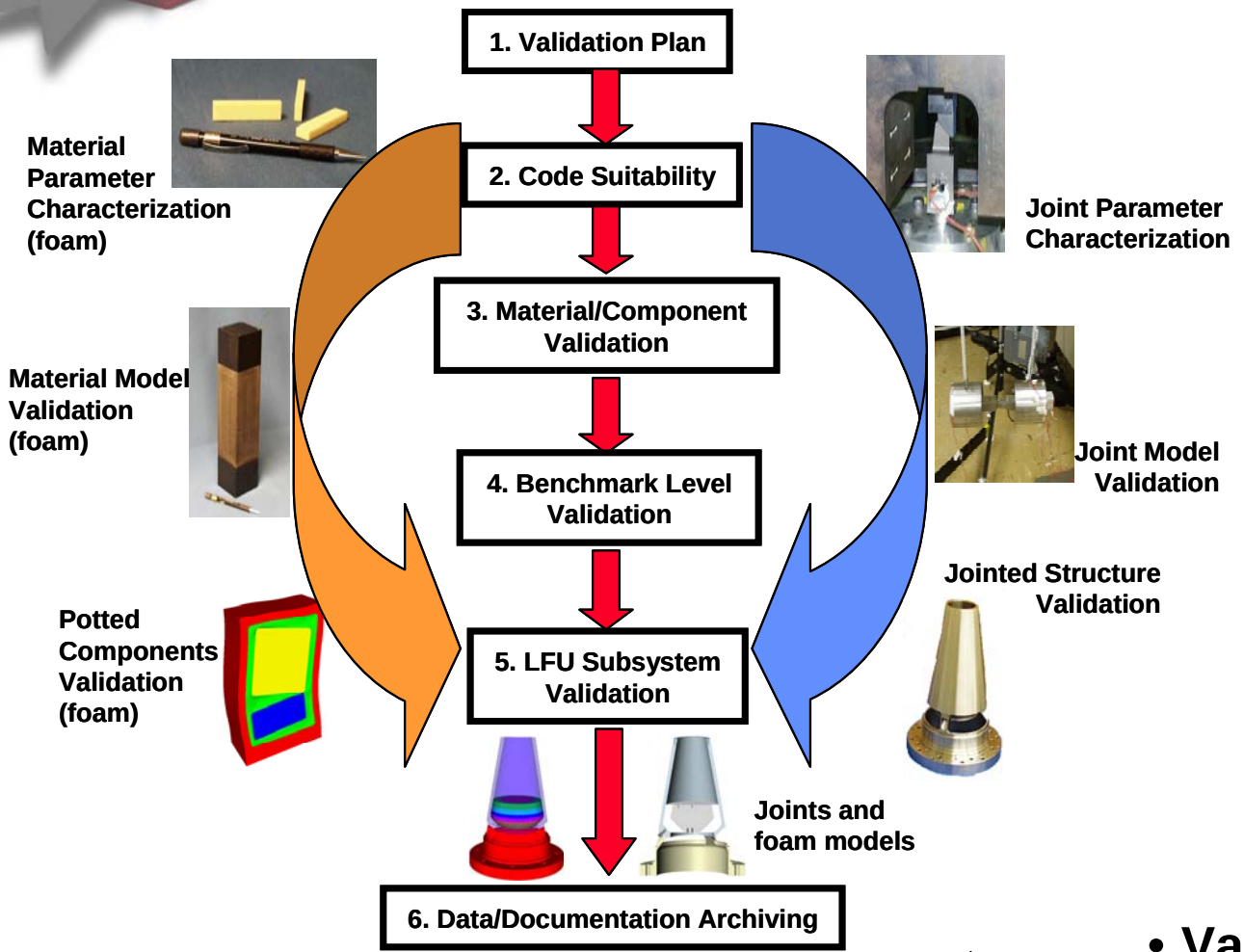


- Calibrate model to blast test using global stiffness and damping parameters: knobs that act as surrogates for missing or unknown physics

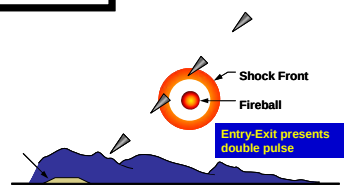


- Use calibrated model to make prediction in tactical environments

Improving Physics Fidelity



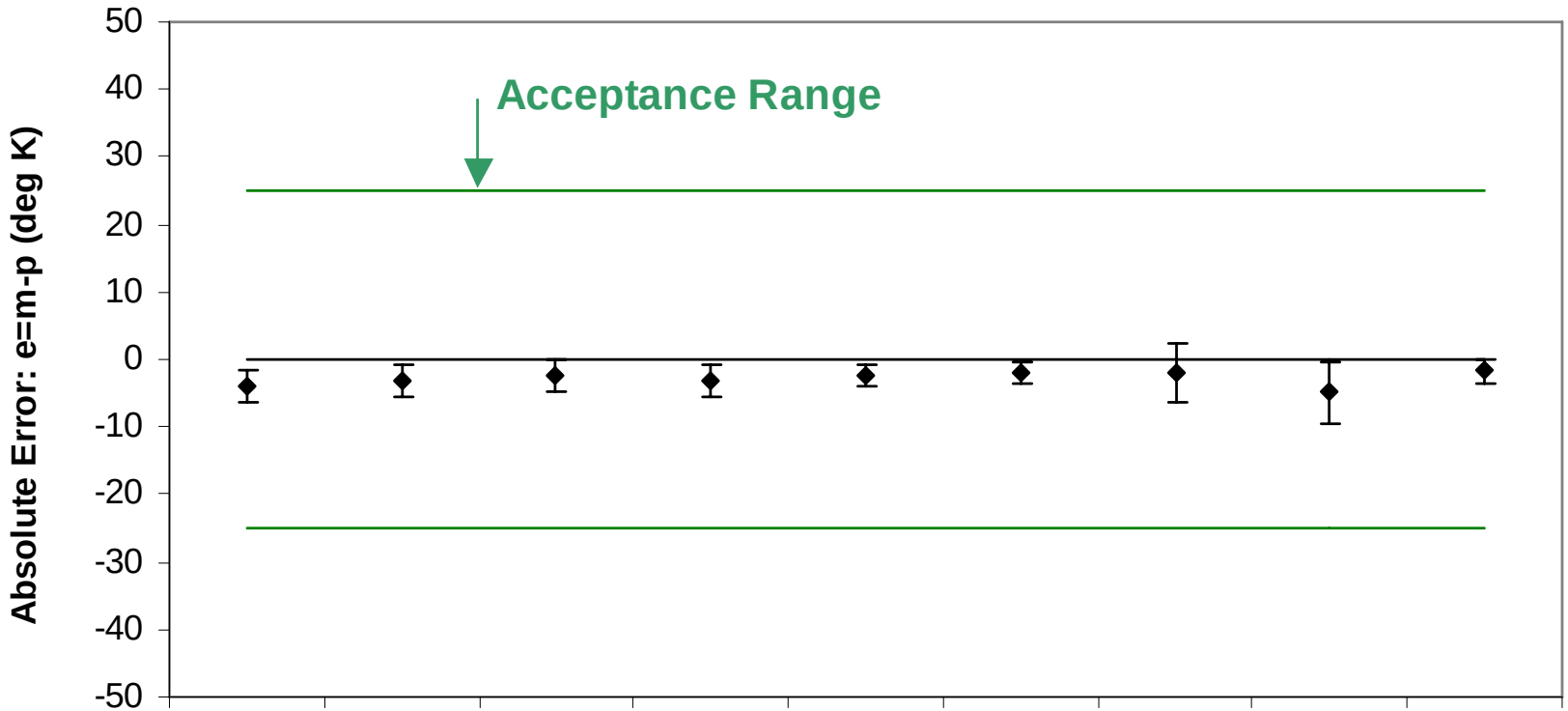
- **Physics-informed models validated against separate effects tests**



- **Validate against blast test and make prediction in tactical environments**

Well Established Physics Fidelity

e~2K for conduction



Various Locations in Test Object

Code Readiness

Hyperlinks

Attributes of Verification

Demonstrating Convergence to Correct Answer for the Intended Application

Solution Verification: Convergence for intended application, *but is it the right answer?*

- *Address adequacy of spatial AND temporal AND other discretizations AND numerical knobs*

Inference → Application

Coupled Multi-Physics Across Codes

Coupled Multi-Physics Within Code

Separate Physics

Issues: non-smooth solutions, contact, constitutive laws, internal constraints, multiscale physics, global/local norms, etc.

Inference

Regression Testing

Code Verification: Convergence to correct answer, *wrong application*

- Eliminate code bugs *AND* inadequate algorithms

Code to Code Comparisons Are a Poor Substitute for Formal Verification

Code Comparison Principle (CCP)

Code 1 = assessed code **Code 2 = benchmark code**

$$\|\text{Code 1} - \text{Truth}\| \le \|\text{Code 1} - \text{Code 2}\| + \|\text{Code 2} - \text{Truth}\|$$

$$\|\text{Code 1} - \text{Code 2}\|$$

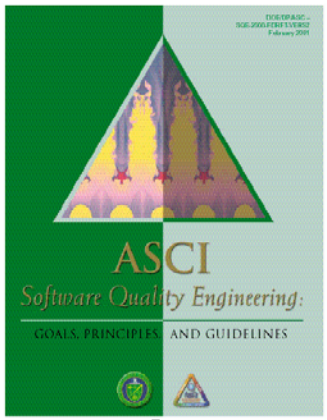
What if this term is not negligible?

- **Could be that Code 1 models are different from Code 2 models**
- **Could be a bug in Code 1 or Code 2**
- **Could be an algorithm flaw in Code 1 or Code 2**
- **Could be that Code 1 or Code 2 model is not converged**

Points to path for better code-to-code comparisons; but if Code 2 is formally verified, why not verify Code 1 to the same verification test suite? And if not, why bother with the code-to-code comparison?

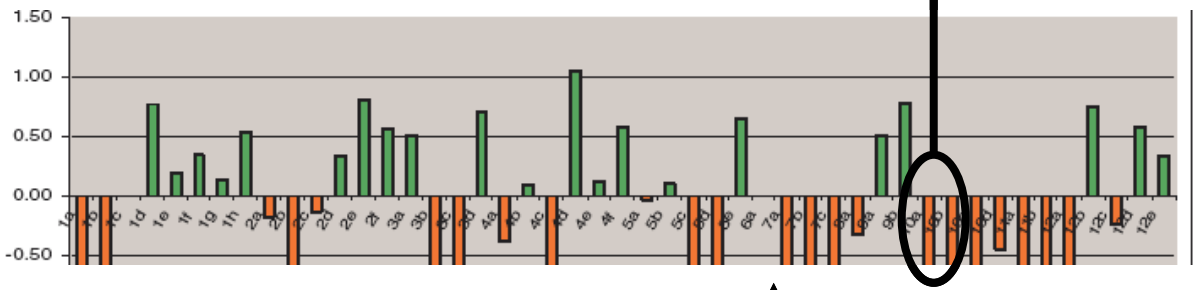
SQE(A): Demonstrated Due Diligence in the Stewardship of Codes

Requirements



SourceForge: Issue Tracking

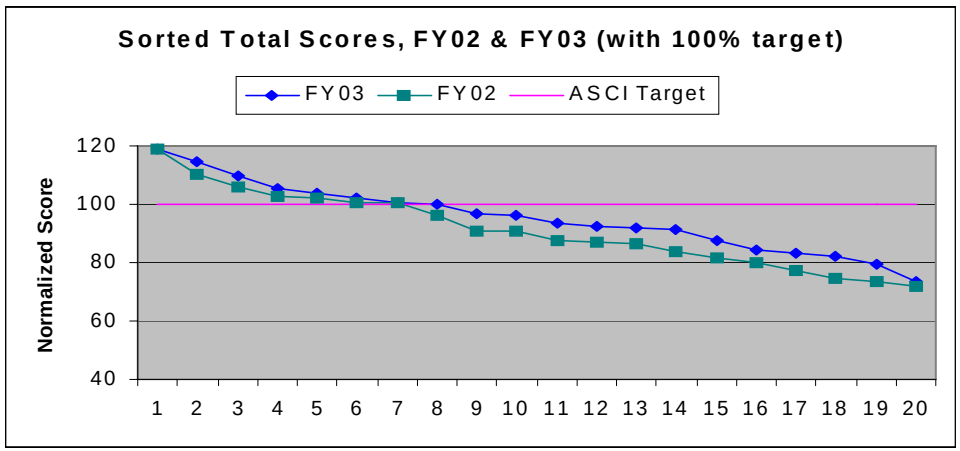
Improvement



FY03 AVG = 95.36

Calore = 96

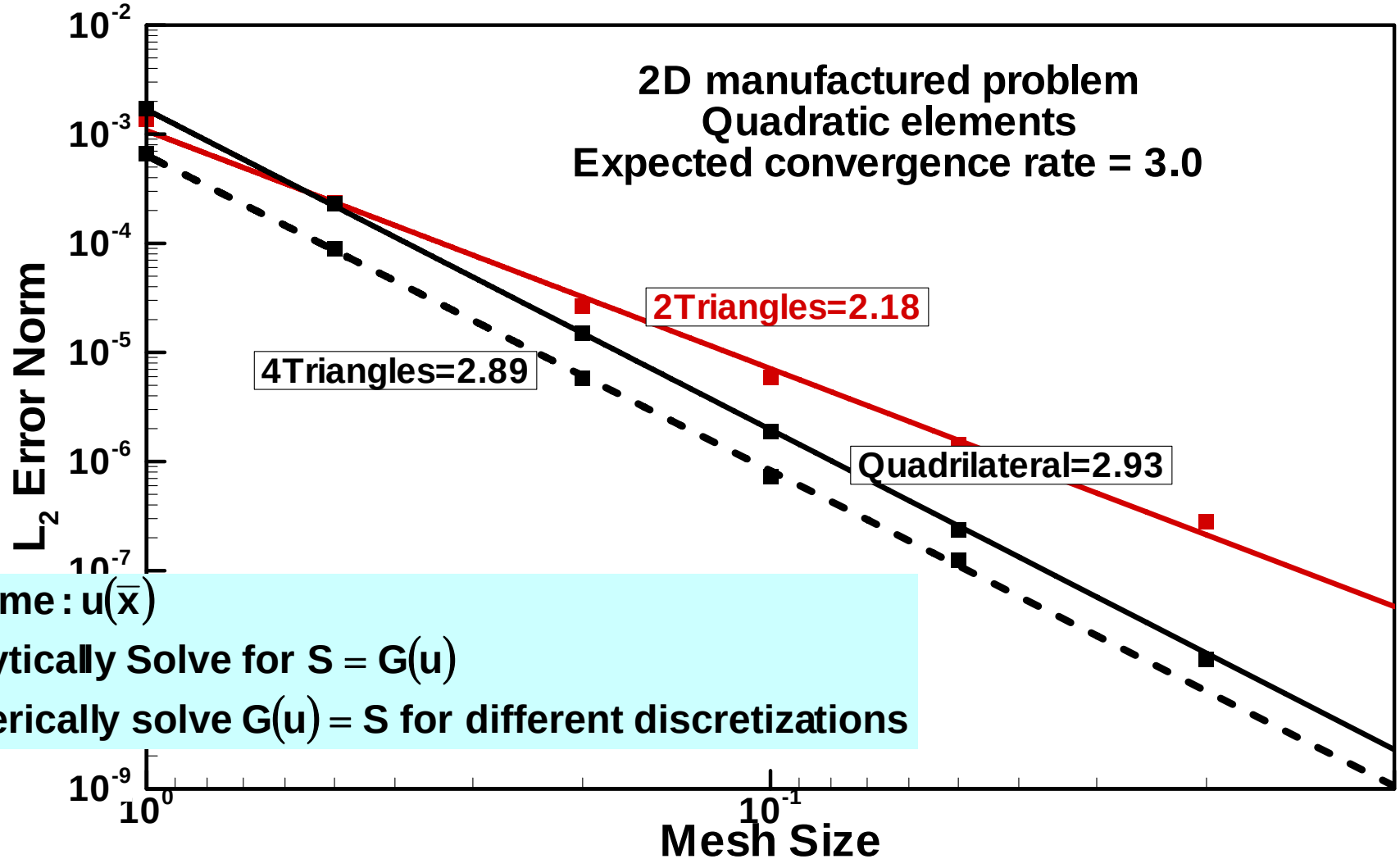
FY02 AVG = 91.06



Assessments

Verification with Manufactured Solution

CEPTRE: Radiation Transport



Assume : $u(\bar{x})$

Analytically Solve for $S = G(u)$

Numerically solve $G(u) = S$ for different discretizations

Measuring Progress in Code Verification Coverage and Interactions

	Features & Capabilities	Unit Tests	Verification Test Suite: VERT 1	VERT 2	VERT 3	Ideal
Code A	FC1		VT1			
	FC2	UT1	VT1			
	FC3	UT2	VT1			
	FC4	UT3	VT1			
	FC5			VT2		
Code B	FC6	UT4		VT2		
	FC7	UT5			VT3	
	FC8	UT6			VT3	
	FC9	UT7			VT3	
	FC10	UT8			VT3	
	Code or Appl Perspective	Line or Cap Coverage	Capability+Interaction Coverage			
		80%		3.22%		

$$f = \frac{\sum_{i=1}^{Nverts} \left(\sum_{r=1}^{nv} {}_{nv}C_r \right)_i}{\sum_{r=1}^{NFC} {}_{NFC}C_r}$$

Solution Verification

Hyperlinks

Attributes of Verification

Demonstrating Convergence to Correct Answer for the Intended Application

Solution Verification: Convergence for intended application, *but is it the right answer?*

- *Address adequacy of spatial AND temporal AND other discretizations AND numerical knobs*

Inference → Application

Coupled Multi-Physics Across Codes

Coupled Multi-Physics Within Code

Separate Physics

Issues: non-smooth solutions, contact, constitutive laws, internal constraints, multiscale physics, global/local norms, etc.

Inference

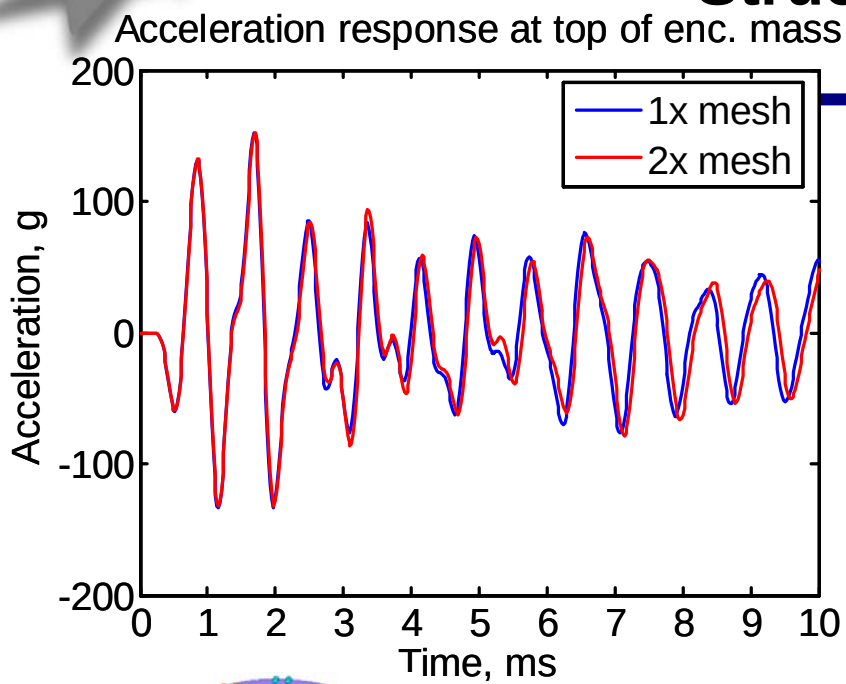
Regression Testing

Code Verification: Convergence to correct answer, *wrong application*

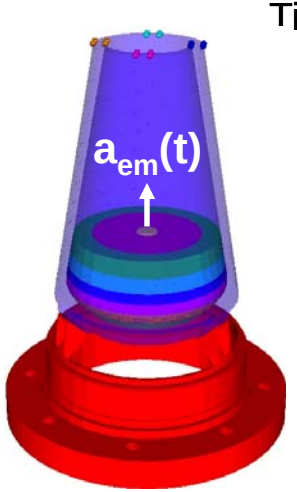
- Eliminate code bugs *AND* inadequate algorithms

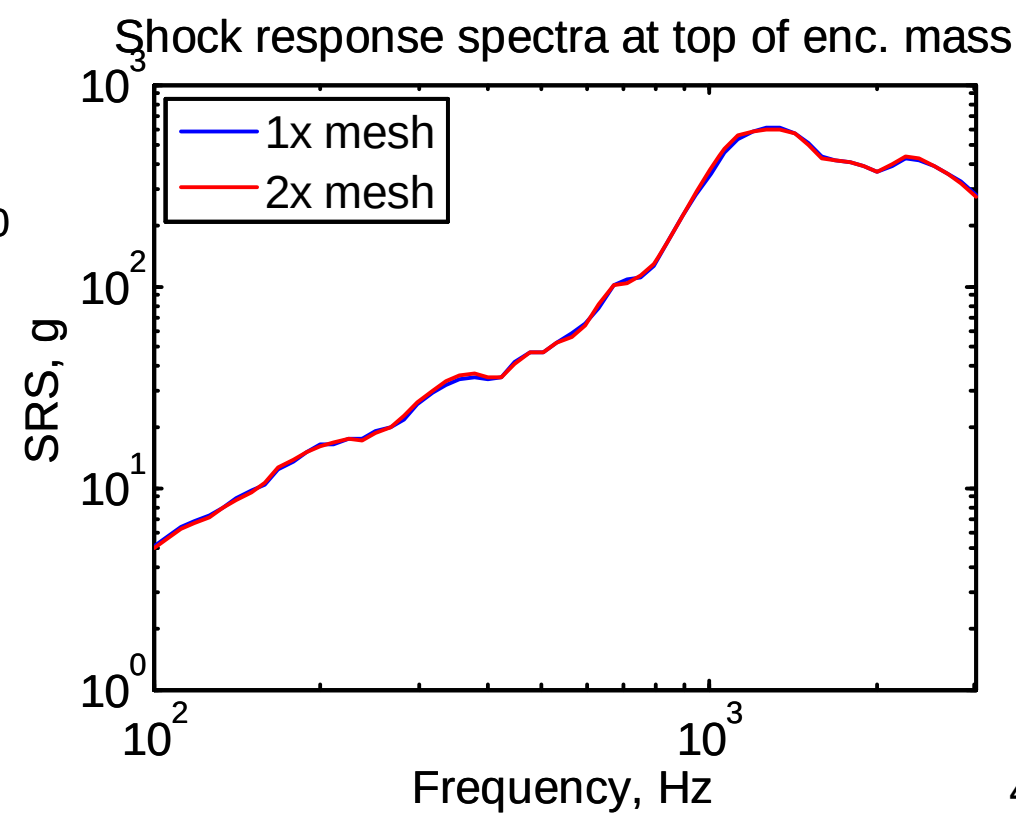
Sensitivity to Mesh Parameters

Structural Dynamics



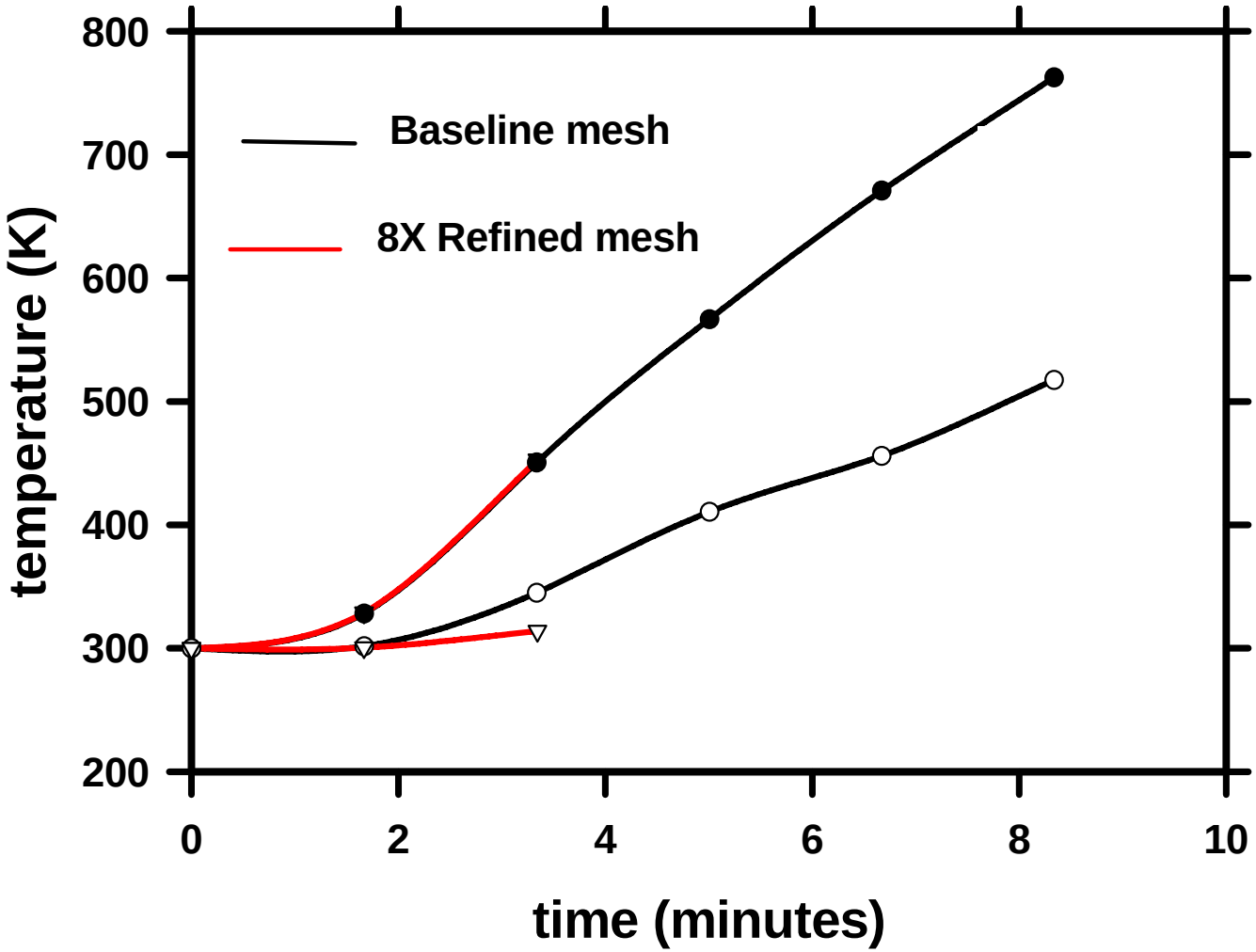
Max. relative error between SRS: +/- 5%



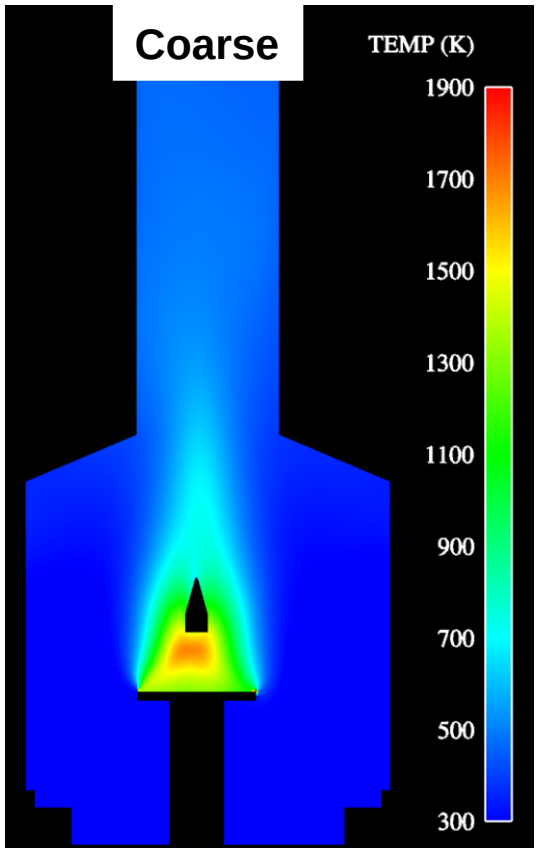


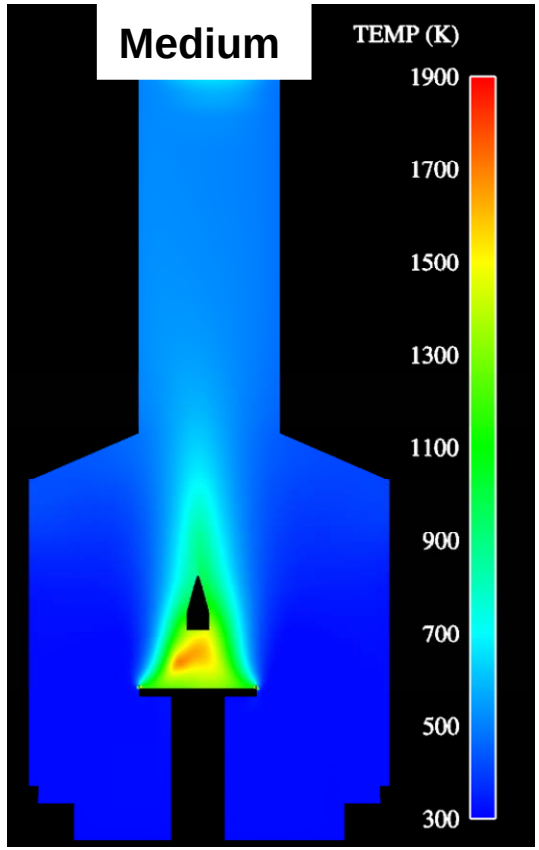
Solution Verification on High Fidelity Models is Hard

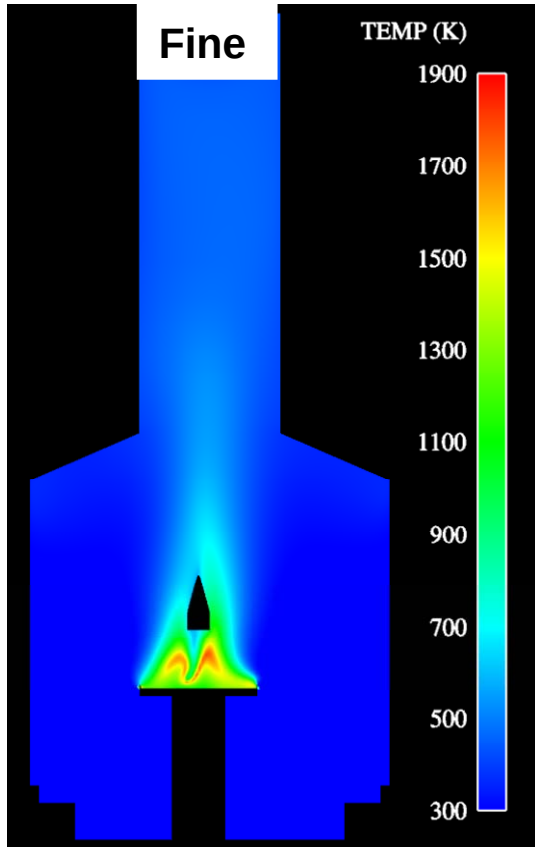
Solution Verification: Is the Discretization Adequate?



Calorimeter Fire
BVG Solutions

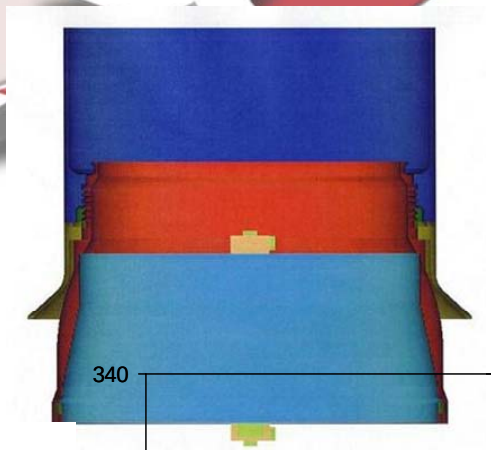


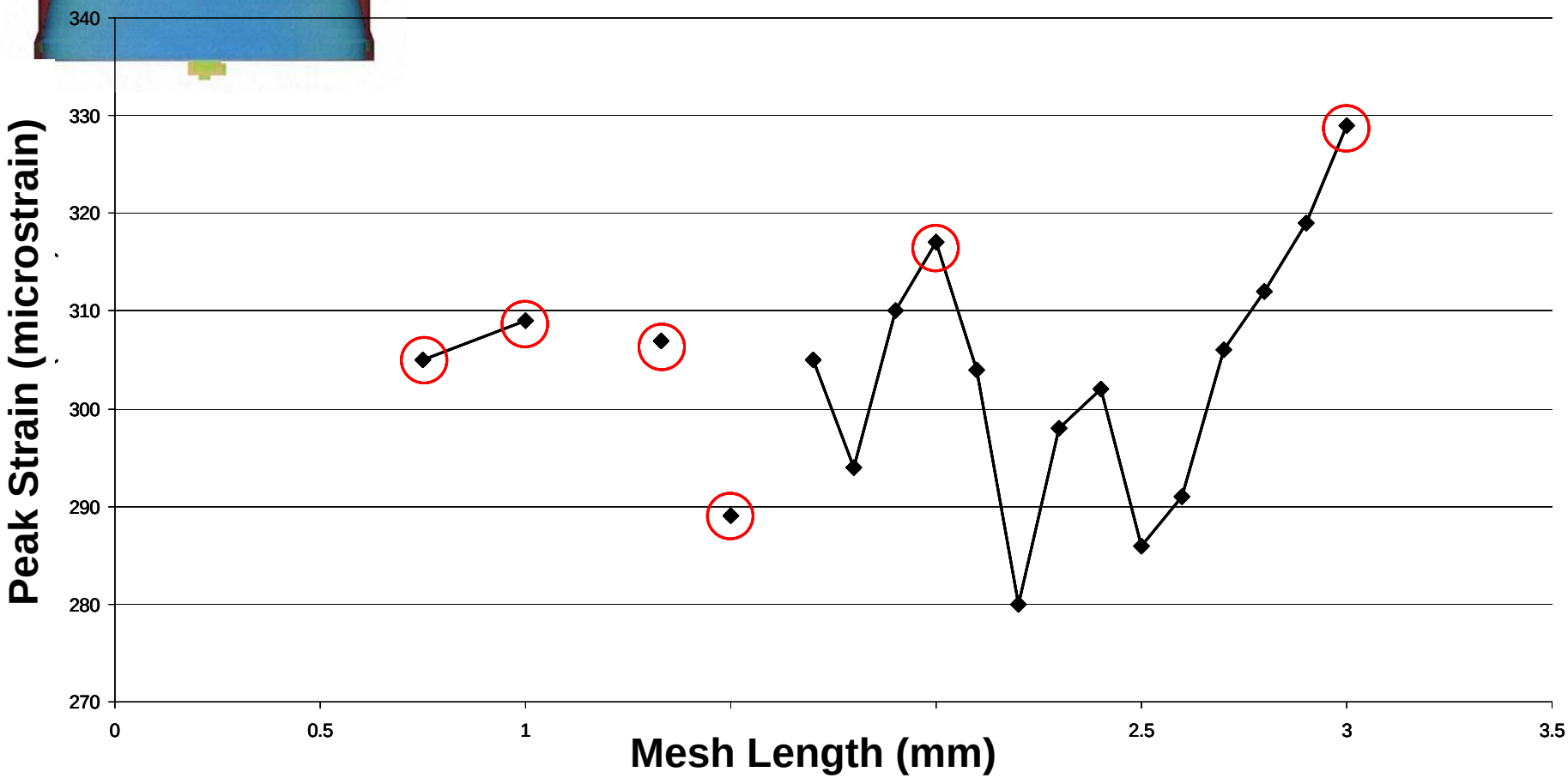




Calculation Verification for a Threaded Assembly

Ryan Maupin, ESA-WR, LANL: IMAC-XXIV 1/31/06

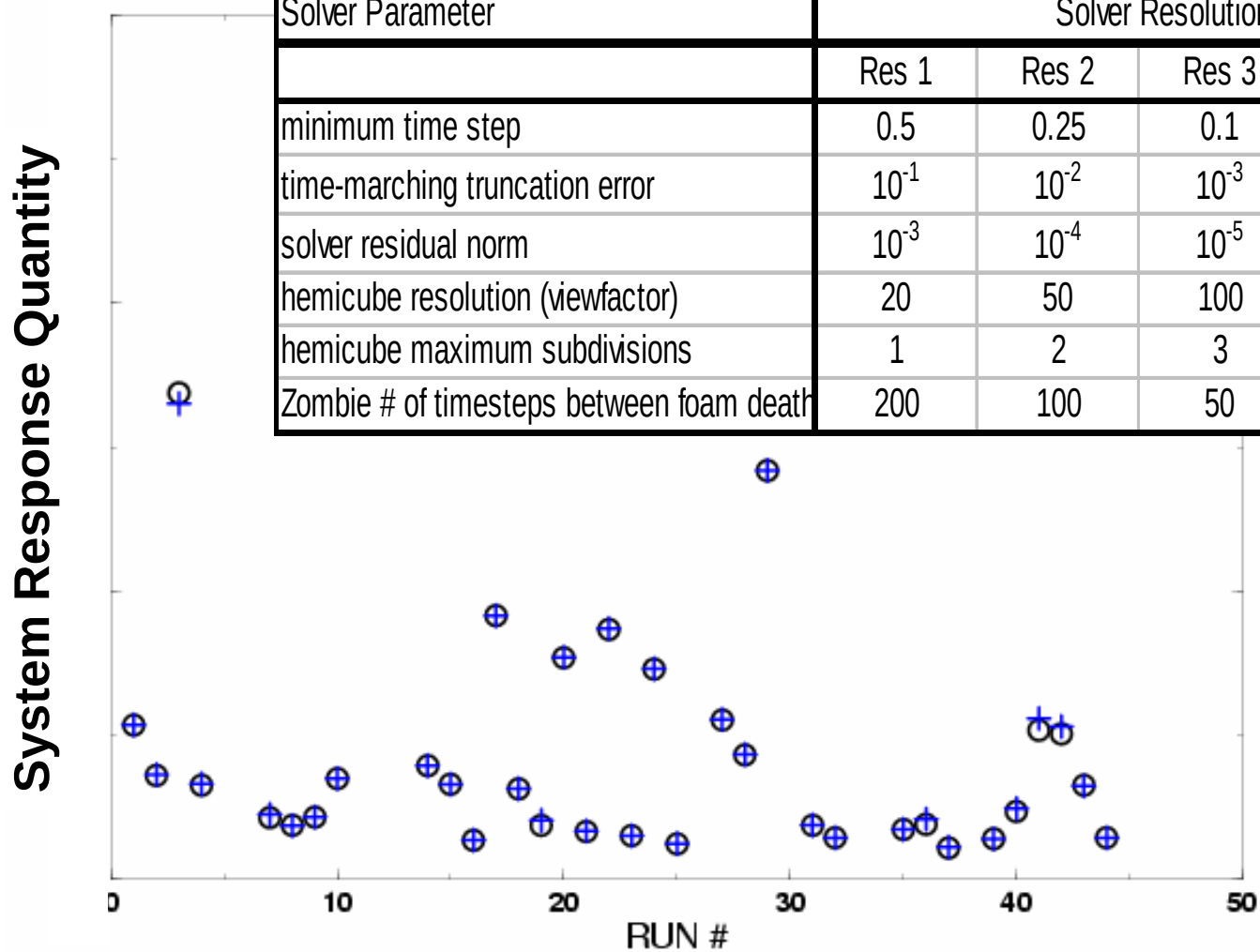




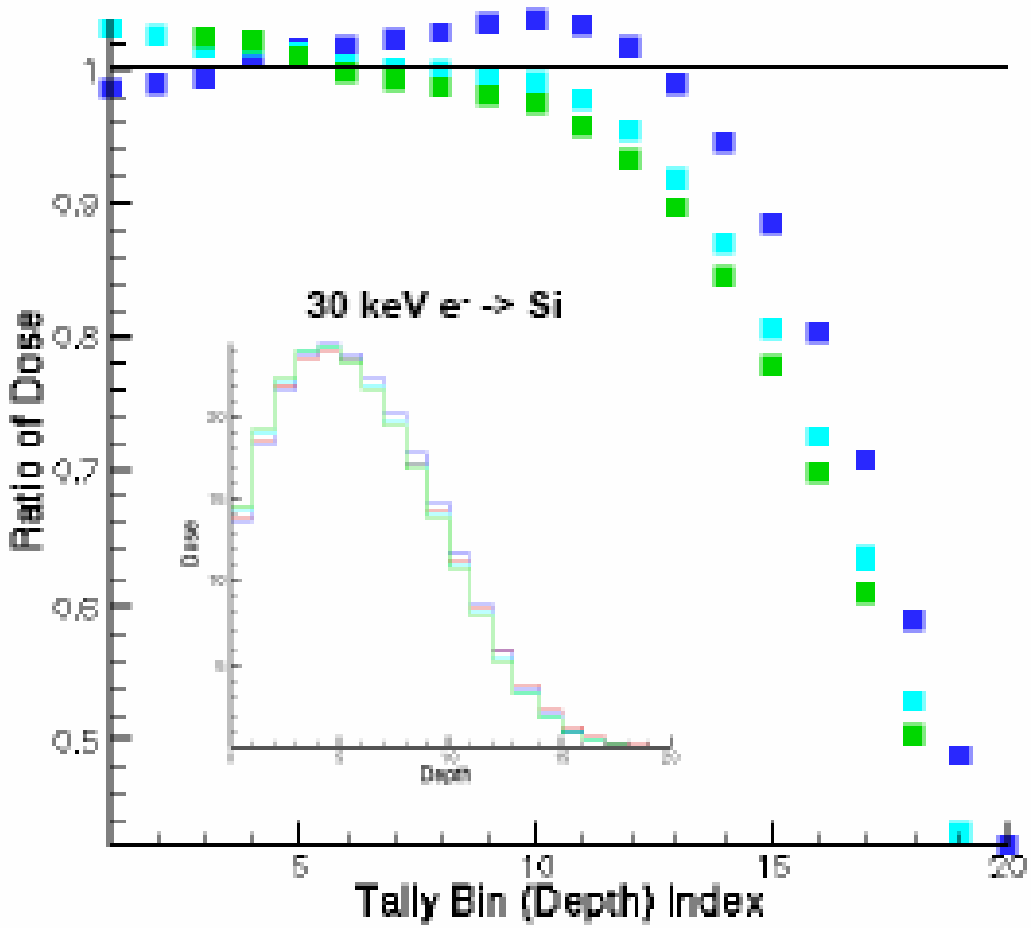
Solver Resolution Over UQ Parameter Space

Solution Verification: Are the solver settings adequate?

Solver Parameter	Solver Resolution Settings				
	Res 1	Res 2	Res 3	Res 4	Res 5
minimum time step	0.5	0.25	0.1	0.05	0.01
time-marching truncation error	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}
solver residual norm	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
hemicube resolution (viewfactor)	20	50	100	200	300
hemicube maximum subdivisions	1	2	3	4	5
Zombie # of timesteps between foam death	200	100	50	20	1

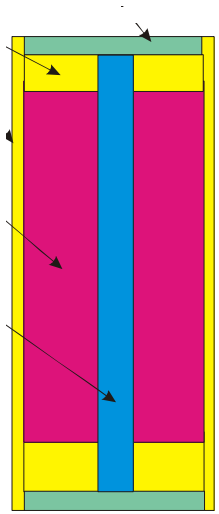


Dose Sensitivity to Electron Boundary Crossing Algorithm

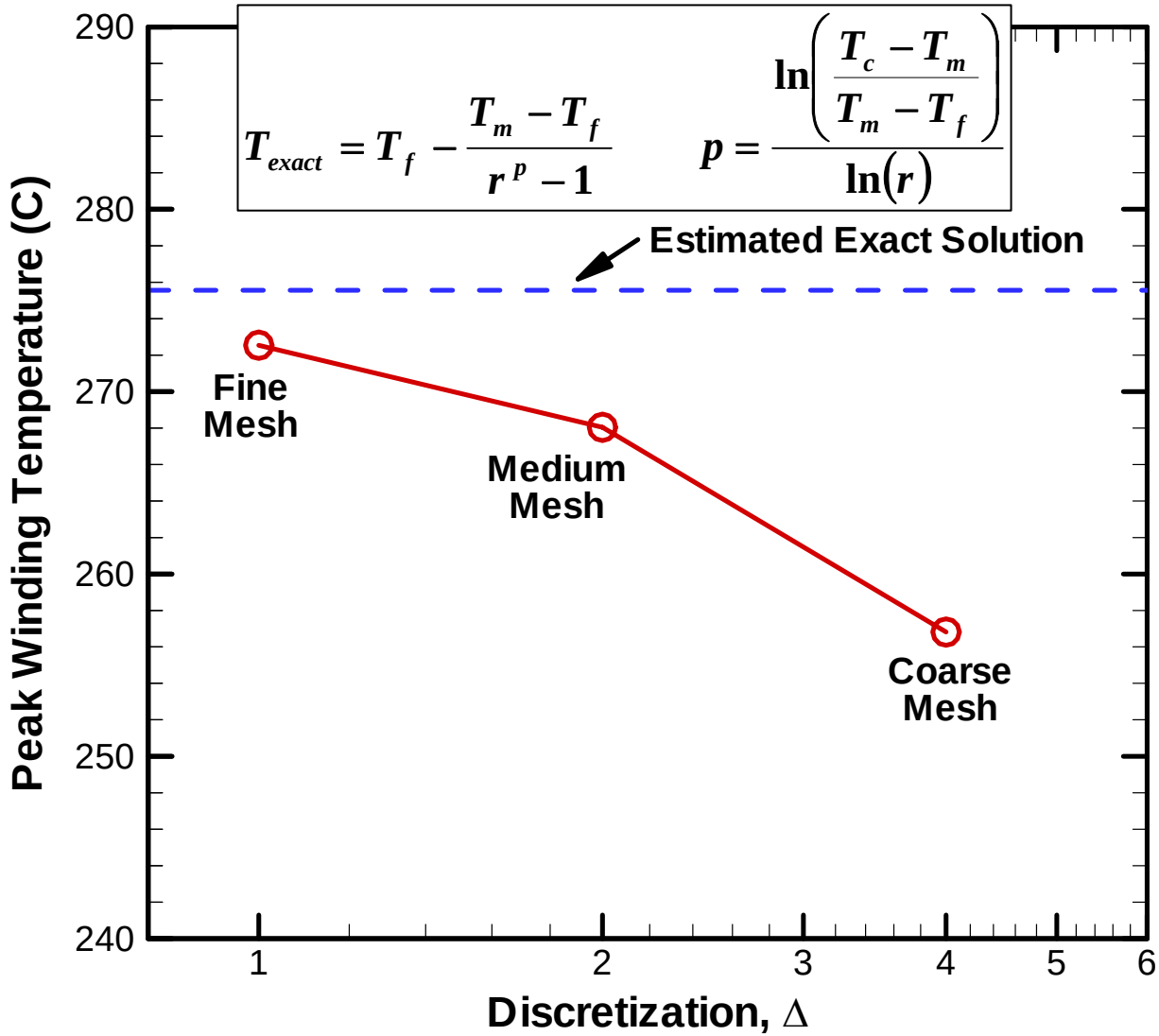


- Evaluation of ITS electron boundary-crossing error: (All with respect to no internal boundaries, default substep size.
- Blue: internal boundaries, default subtep size
- Cyan: Internal boundaries half-default substep size
- Green: Internal boundaries quarter-default substep size

Numerical Errors Pollute Validation Assessments



$$T_{exact} = T_f - \frac{T_m - T_f}{r^p - 1} \qquad p = \frac{\ln\left(\frac{T_c - T_m}{T_m - T_f}\right)}{\ln(r)}$$



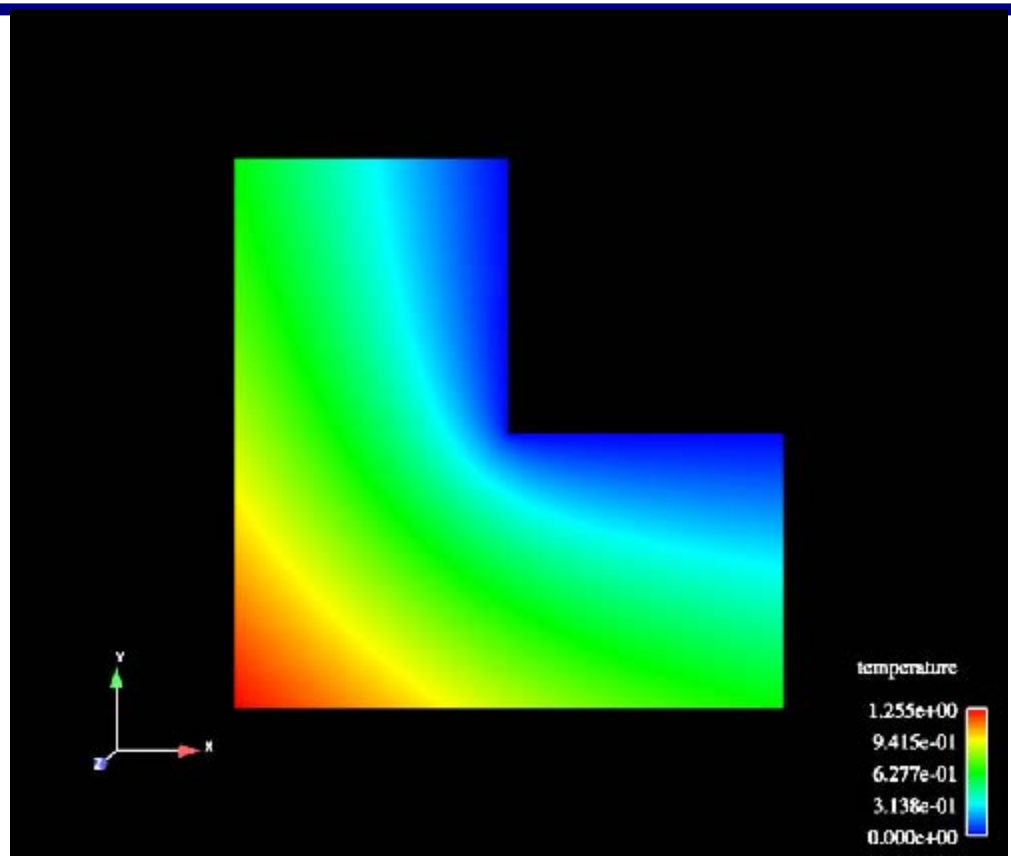
Verification of Error Estimator and Adaptive Algorithm

- 2D Exact Solution:

$$u = r^{2/3} \sin\left(\frac{2}{3}\theta\right)$$

- Linear elements
- ZZ error estimator

- Feedback adaptive algorithm:



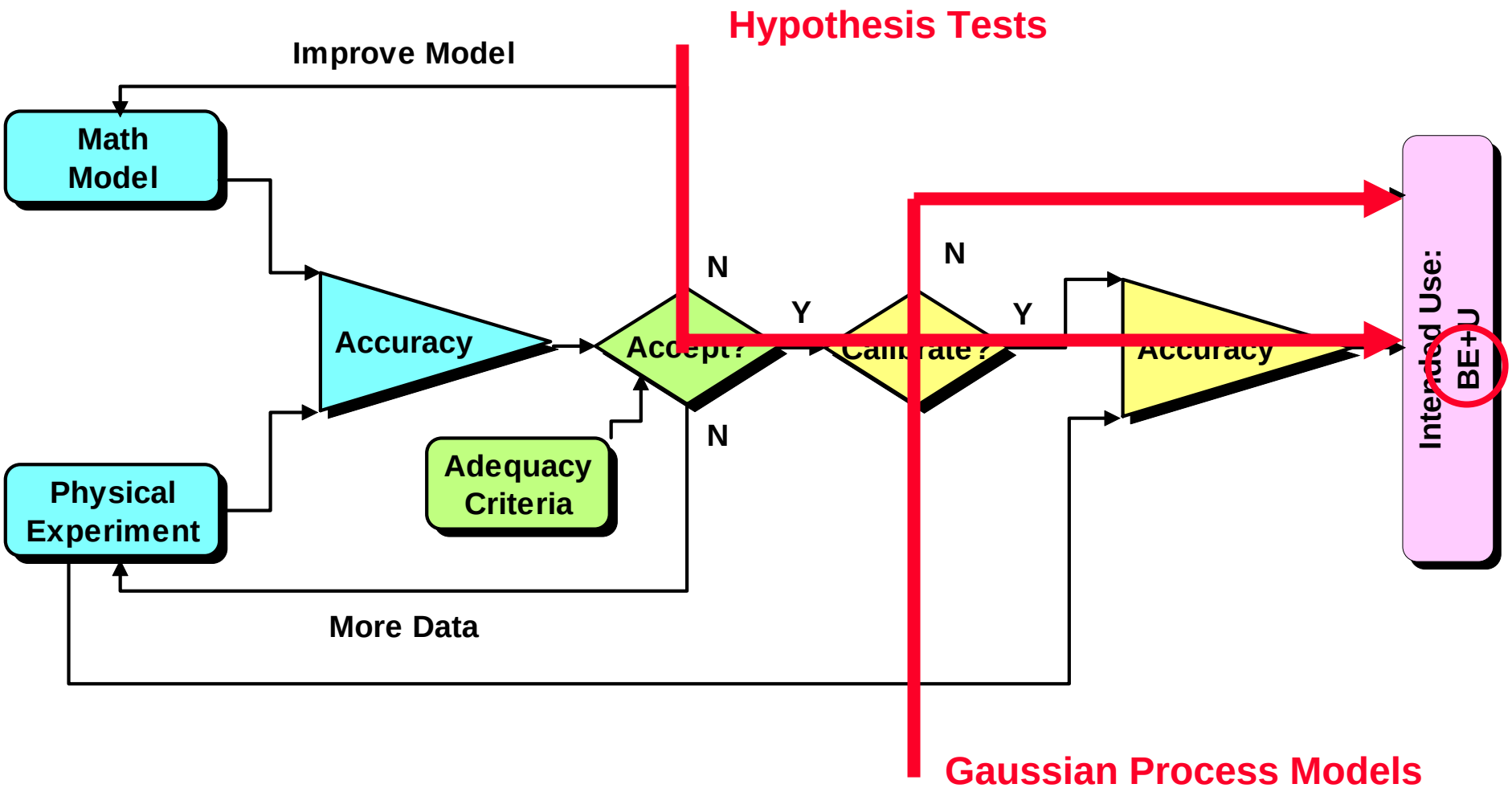
$$\text{if } \left(||e^*||_{H_1(\omega_j)} > 0.995 \max_{1 \le i \le N_\omega} ||e^*||_{H_1(\omega_i)} \right) \text{ then refine } \omega_j$$

Validation

Hyperlinks

Validation is Assessment Supporting BE+U

Calibration is not Validation

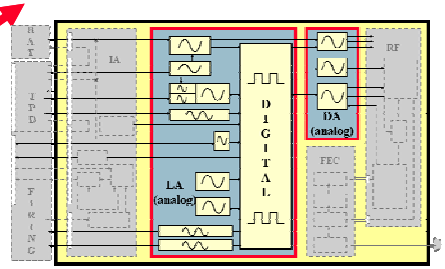


Science-Based Validation Experiments

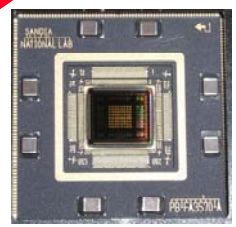
Validation: Are You Solving the Right Equations?

Hierarchal Validation: Right answer for the right reason

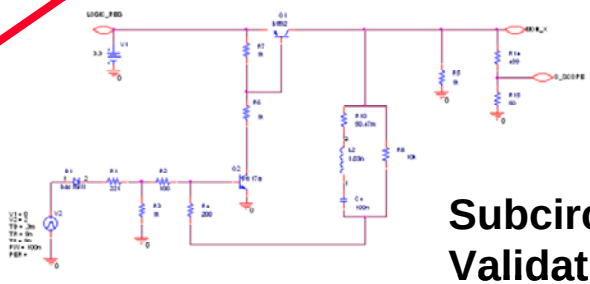
Increasing complexity, Decreasing number of tests



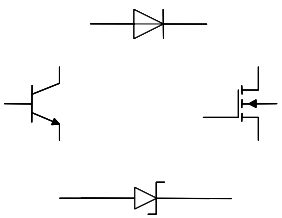
System-Level Circuit Validation



Single ASIC Validation



Subcircuit Validation



Single Device Characterization and Validation

- **Application relevant parameter space**
- **Formal DOE and replicate tests**
- **Attention to diagnostic bias and precision**

Science-Based Validation Experiments

Validation: Are You Solving the Right Equations?

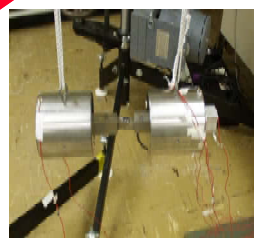
Hierarchal Validation: Right answer for the right reason

Increasing complexity, Decreasing number of tests

Full System Test

Mockup with jointed structure and foam embedded object

Jointed structure validation



Single joint validation

Joint parameter characterization

- **Application relevant parameter space**
- **Formal DOE and replicate tests**
- **Attention to diagnostic bias and precision**

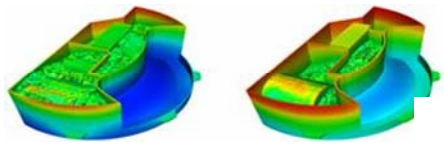
Science-Based Validation Experiments

Validation: Are You Solving the Right Equations?

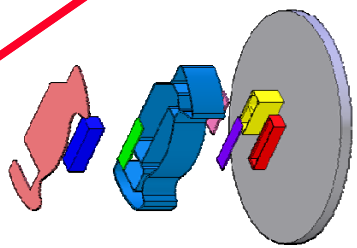
Hierarchal Validation: Right answer for the right reason

Increasing complexity, Decreasing number of tests

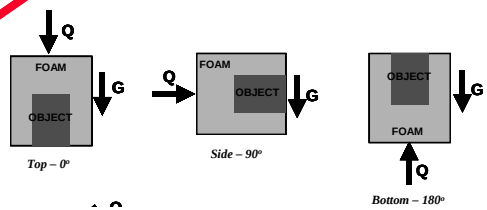
Full System Test



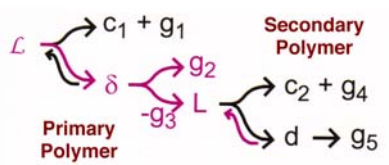
Validation Real Sub-systems



Validation with mockups



Foam recession

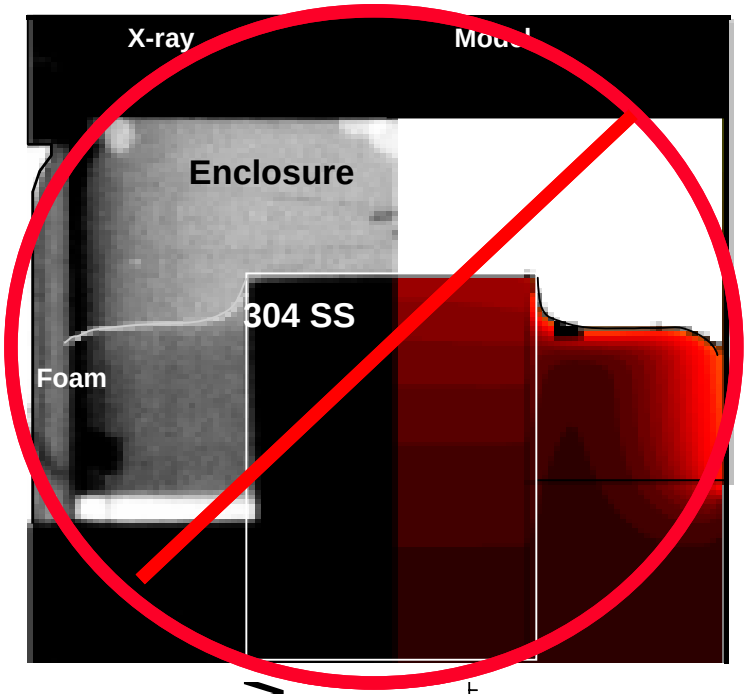


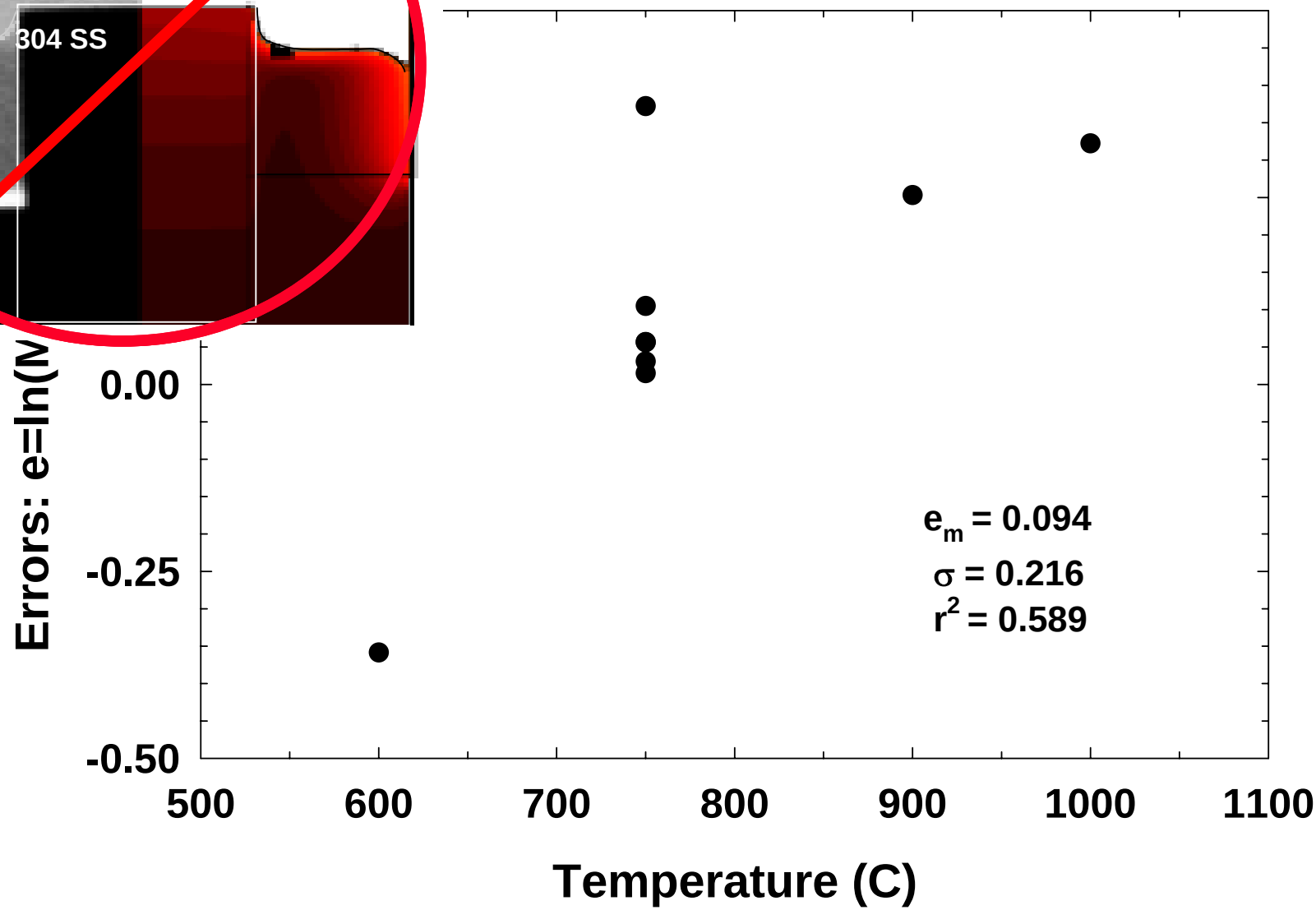
Chemistry characterization/validation

- **Application relevant parameter space**
- **Formal DOE and replicate tests**
- **Attention to diagnostic bias and precision**

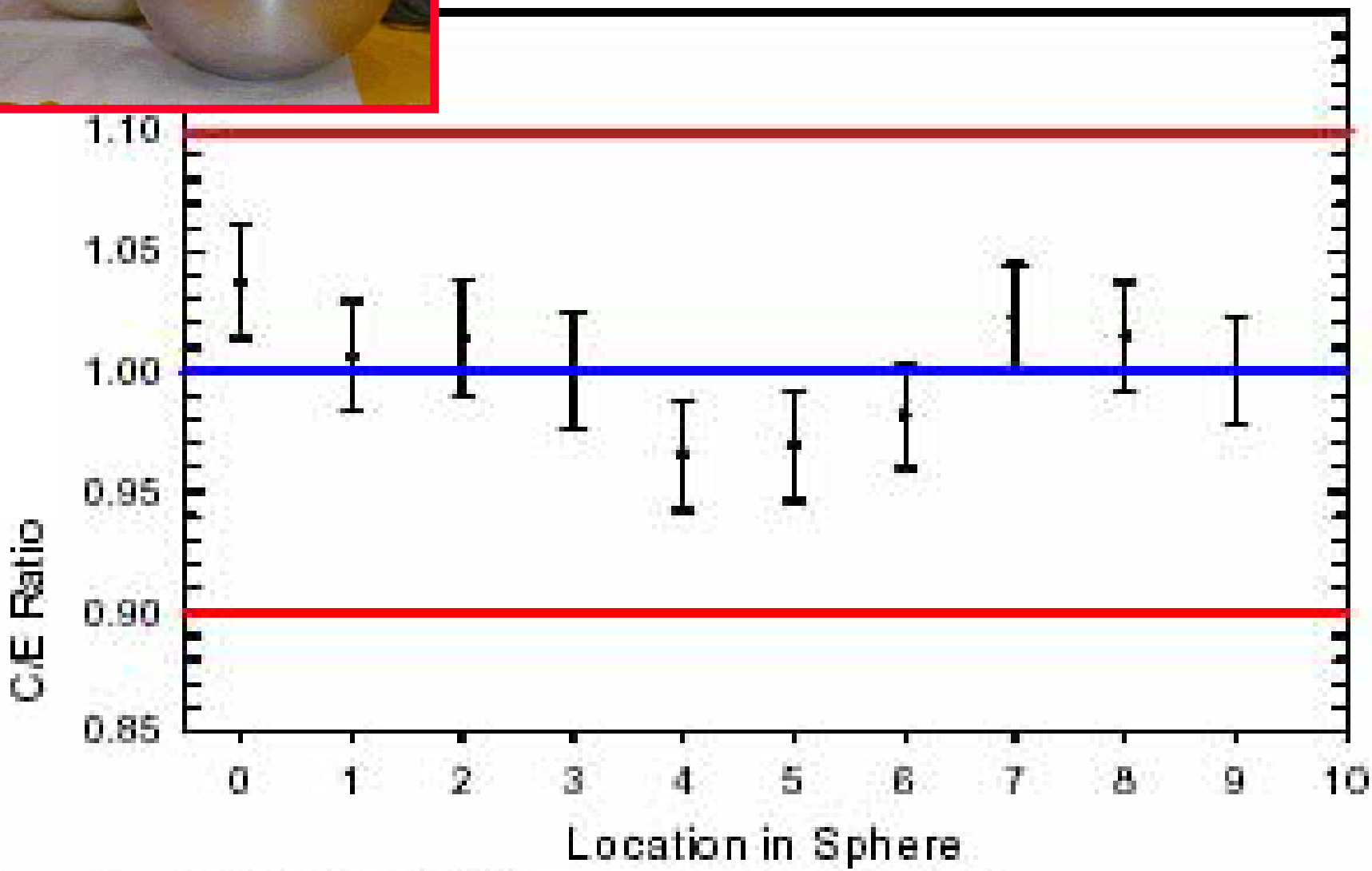
Validation is Statistical

Vugraph Norms Are Not Adequate





Neutron Attenuation in Test Objects

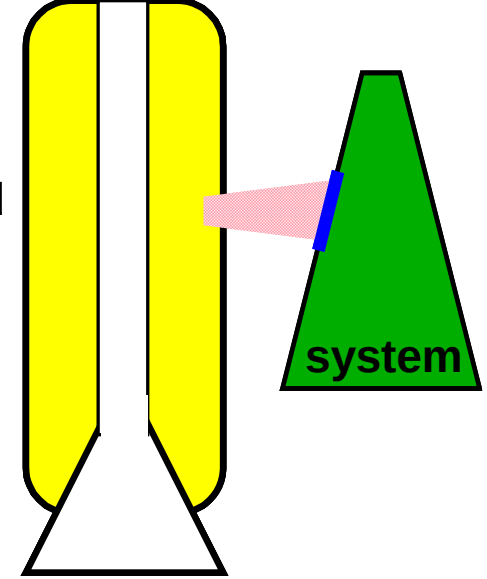


QMU and Sensitivities

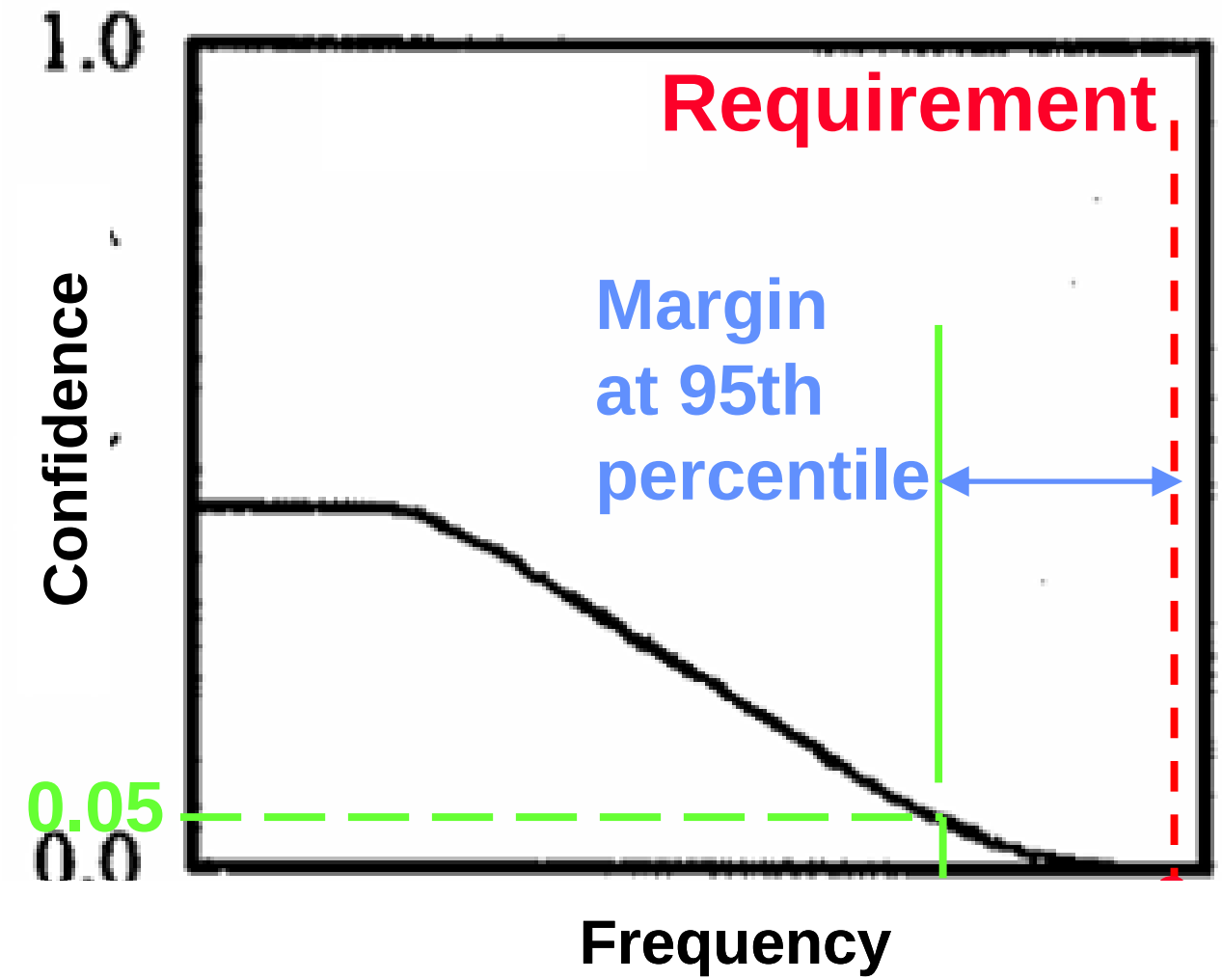
Hyperlinks

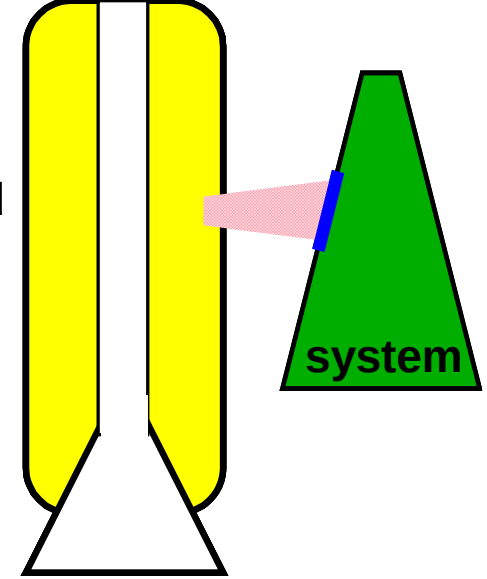
Aleatory and Epistemic Uncertainties

- **Aleatory uncertainty: Inherent randomness in behavior of system under study (frequency interpretation)**
  - **Alternatives: Variability, stochastic uncertainty, irreducible uncertainty, type A uncertainty**
  - **Examples: component failures or material properties derived from statistically significant testing under conditions relevant to intended application**
- **Epistemic uncertainty: Lack of knowledge about appropriate value to use for a quantity that is assumed to have a fixed value in the context of a specific analysis (confidence or belief interpretation)**
  - **Alternatives: state of knowledge uncertainty, subjective uncertainty, reducible uncertainty, type B uncertainty**
  - **Examples: representative scenarios, unknown parameters in frequency distributions, parameters or models with defensible bounds but no sense of frequency**



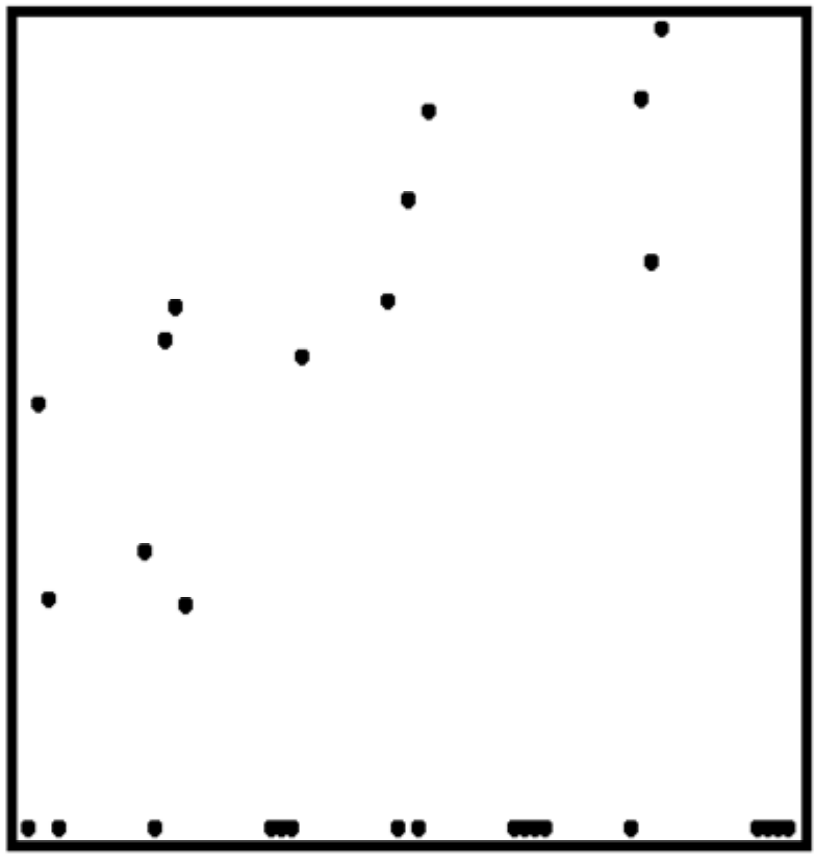
Quantified Margins and Uncertainties

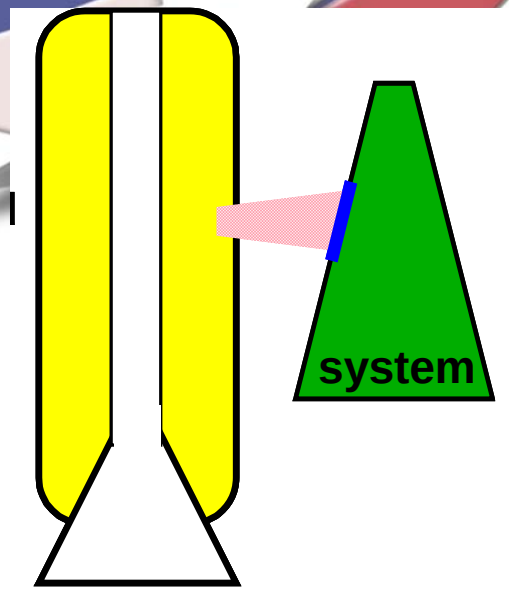




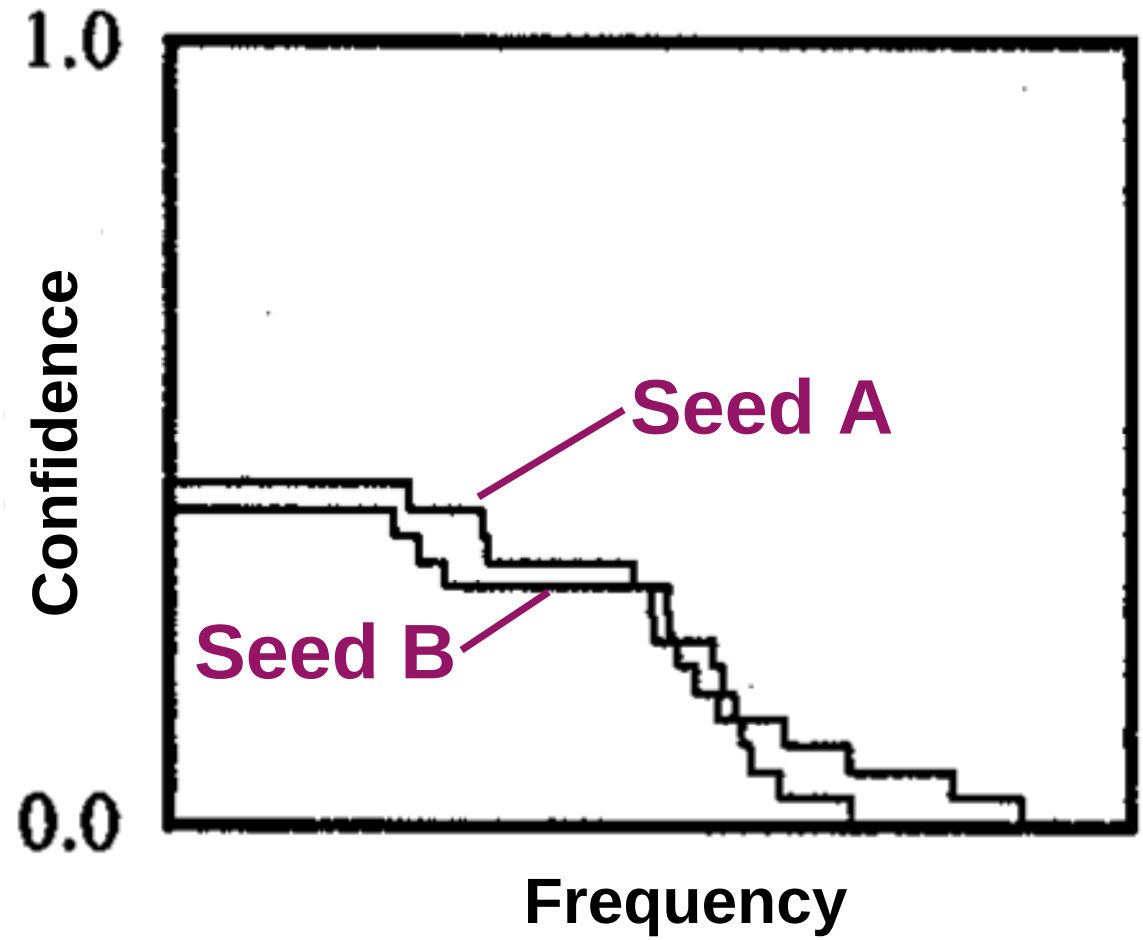
Sensitivity Analysis

$$\mathrm{SCorr} = 0.809$$





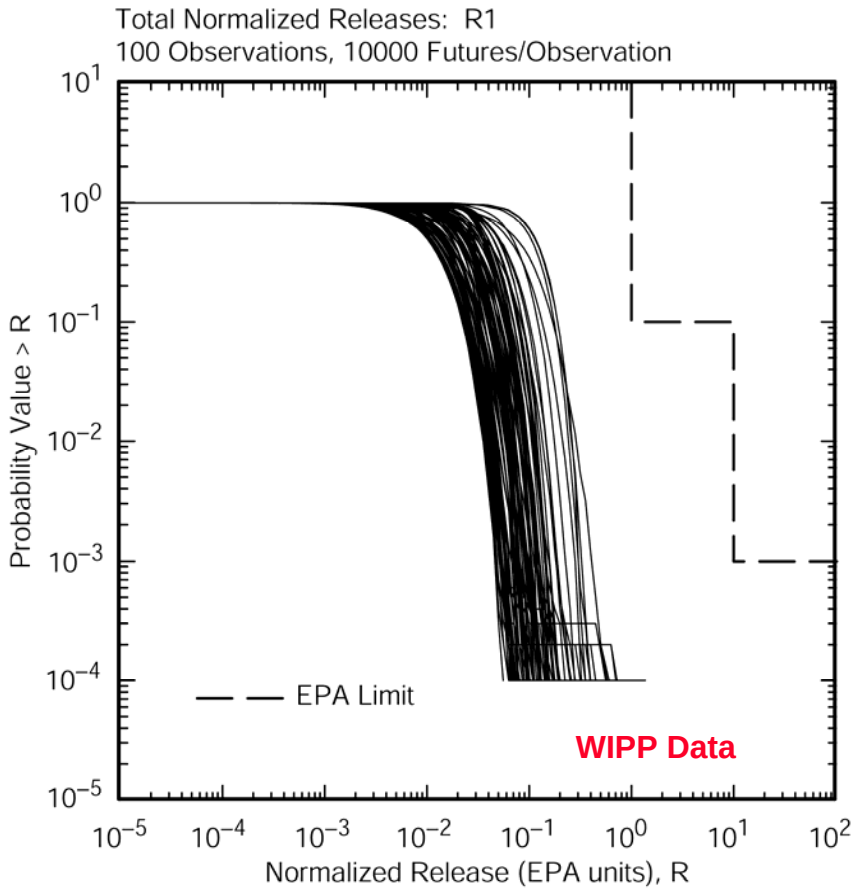
UQ Solution Verification

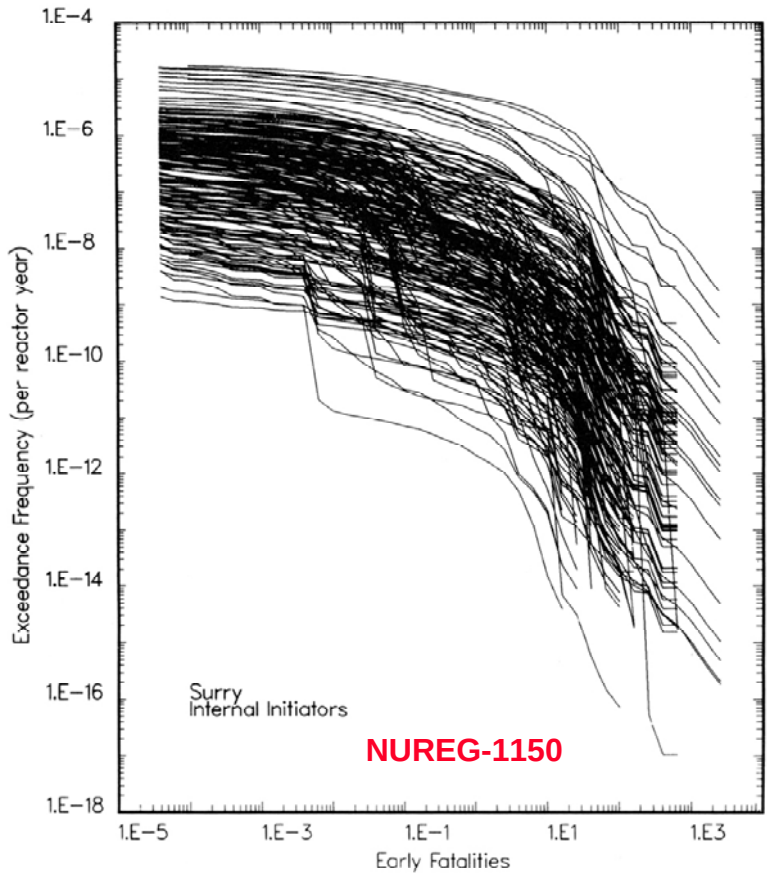


Seed Effects from limited sampling

WIPP and NUREG-1150 Precedents

High Consequence Regulatory Issues in the National Interest Addressed Primary Through Modeling and Simulation





Lessons Learned: (1) Seek BE + Uncertainty

(2) It takes more than one shot to get it right

Infrastructure for ASC-Scale UQ Analyses

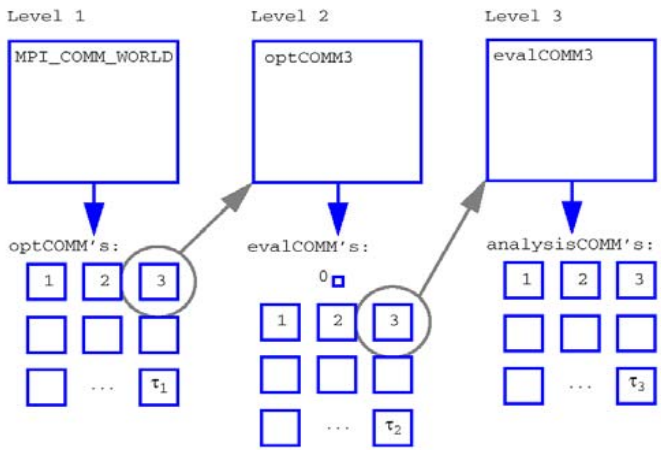
DAKOTA: A framework for optimization/UQ in the ASCI environment

DAKOTA Tools:

- DDACE
- Optimization
- Design of Experiments
- Response Surface

UQ Tools:

- Sampling Tools DMC, LHS
- Reliability Methods FORM, SORM, AMV, AMV+
- Stochastic FEM
- Function Cost Calculator

Optimization Under Variability

1. ***Algorithmic coarse-grained parallelism:*** **independent concurrent fn. evaluations**
2. ***Algorithmic fine-grained parallelism:*** **parallel computation of internal linear algebra**
3. ***Function evaluation coarse-grained parallelism:*** **concurrent execution of separable simulations within a fn. eval. (e.g., multiple loading cases)**
4. ***Function evaluation fine-grained parallelism:*** **parallelization of the solution steps within a single analysis code (e.g., SALINAS, MPSalsa)**