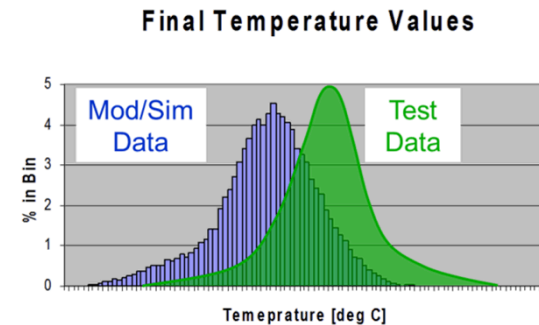
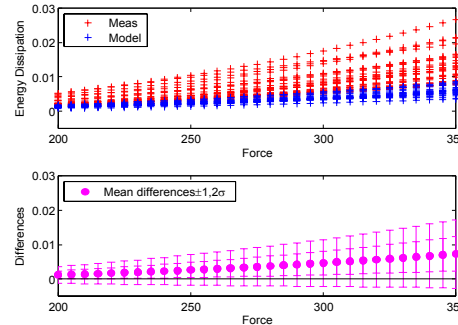
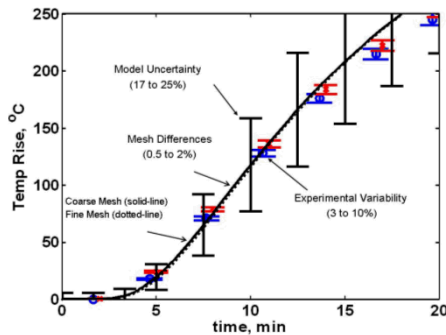
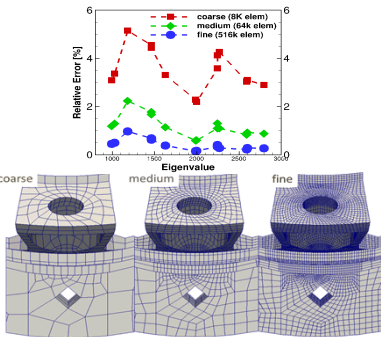


Exceptional service in the national interest



V&V/UQ Activities to Help Support Computational Simulation Credibility



Dr. Walt Witkowski
 Sandia National Laboratories
 Validation, Verification, Uncertainty Quantification
 and Credibility Processes Dept.
wrwitko@sandia.gov

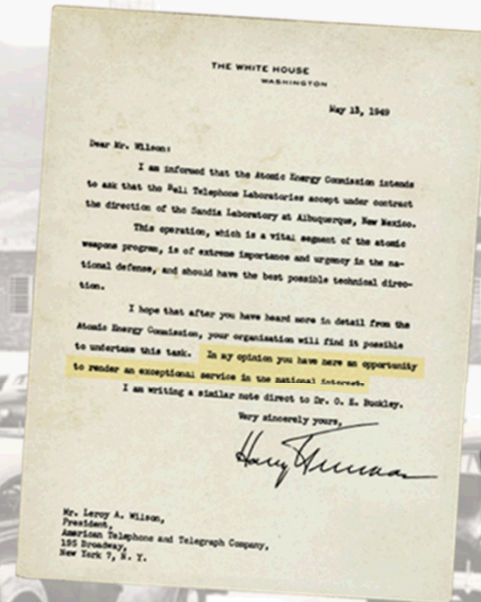
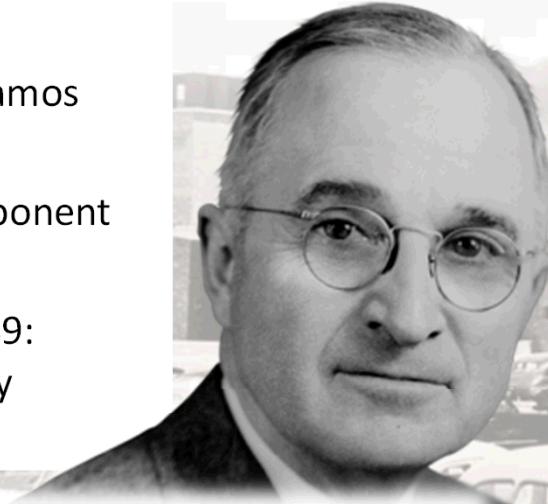


Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. SAND NO. 2011-XXXXP

Exceptional service in the national interest

Then....

- July 1945: Los Alamos creates Z Division
- Nonnuclear component engineering
- November 1, 1949: Sandia Laboratory established



Now....

- Core mission in Nuclear Weapons
- Addressing the nation's most challenging National Security problems
 - USS Iowa (1989)
 - TWA Flight 800 Accident (1997)
 - Post 9/11 Vulnerability Studies (2001)
 - Columbia Space Shuttle Accident (2003)
 - I-35W bridge collapse in Minneapolis (2007)
 - BP Deepwater Horizon Oil Spill Accident (2010)
 - Aircraft Vulnerability (2013)
 - Waste Isolation Plant Leak (2014)

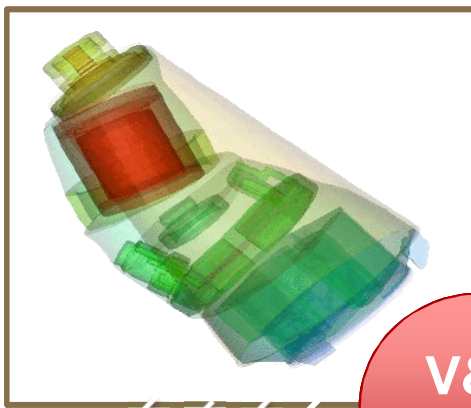


Environmental Tests and Computational Simulation

Capabilities Required to Design, and Qualify the Stockpile

Integrated theory, computational simulation and experimental discovery/validation across length and time scales is critical to develop the technical basis for complex systems.

Engineering
Analysis

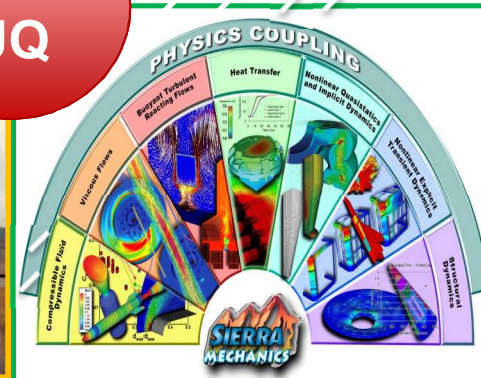


Engineering
Science Physical
Phenomena

Environmental
Simulation &
Test



V&V
UQ



Computational
Simulation
Technology

- SNL nuclear weapon mission drivers:
 - Several systems including 2 current LEPs
 - Support annual assessment & certification that all weapon types are safe, secure & reliable
- Changing paradigm from testing dominated approach
- Few/no tests at the full system level;
few/some/no tests at subsystem/component level:
 - not allowed, and/or (radiation effects tests)
 - too expensive, and/or (crash impact tests)
 - too environmentally unfriendly, and/or (fuel/propellant fire tests)
 - too few units available (annual surveillance)
 - Or ...just cannot perform the test at the application specs

Are we ready for a challenge?

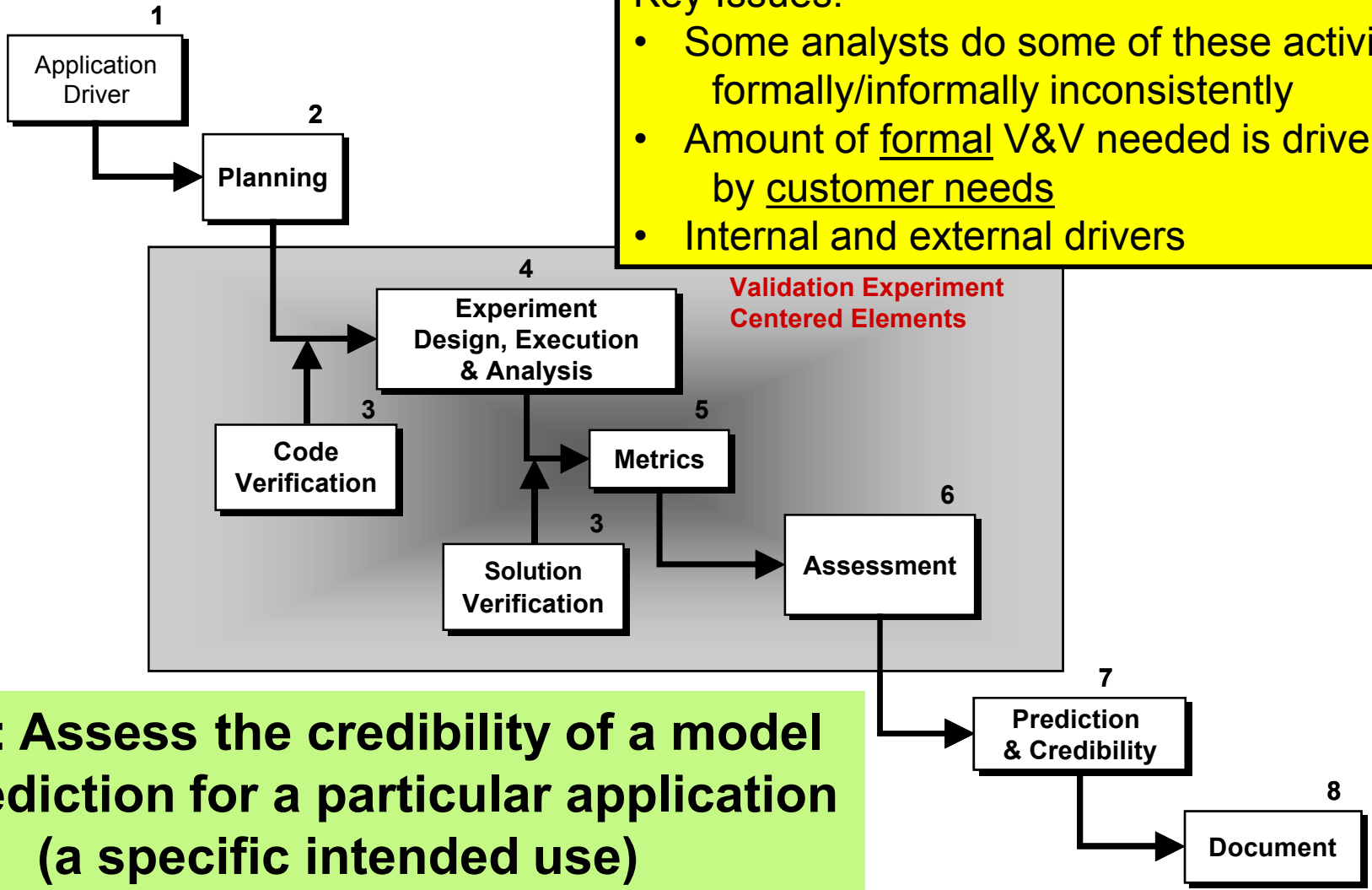
- If called upon, how do we defend our predictions?
- What evidence is necessary to support our claims?
- It depends on what questions will be answered by the decision maker using the simulation results and testing data **AND** the associated risk with the decision being wrong



Sandia V&V Process and Activities

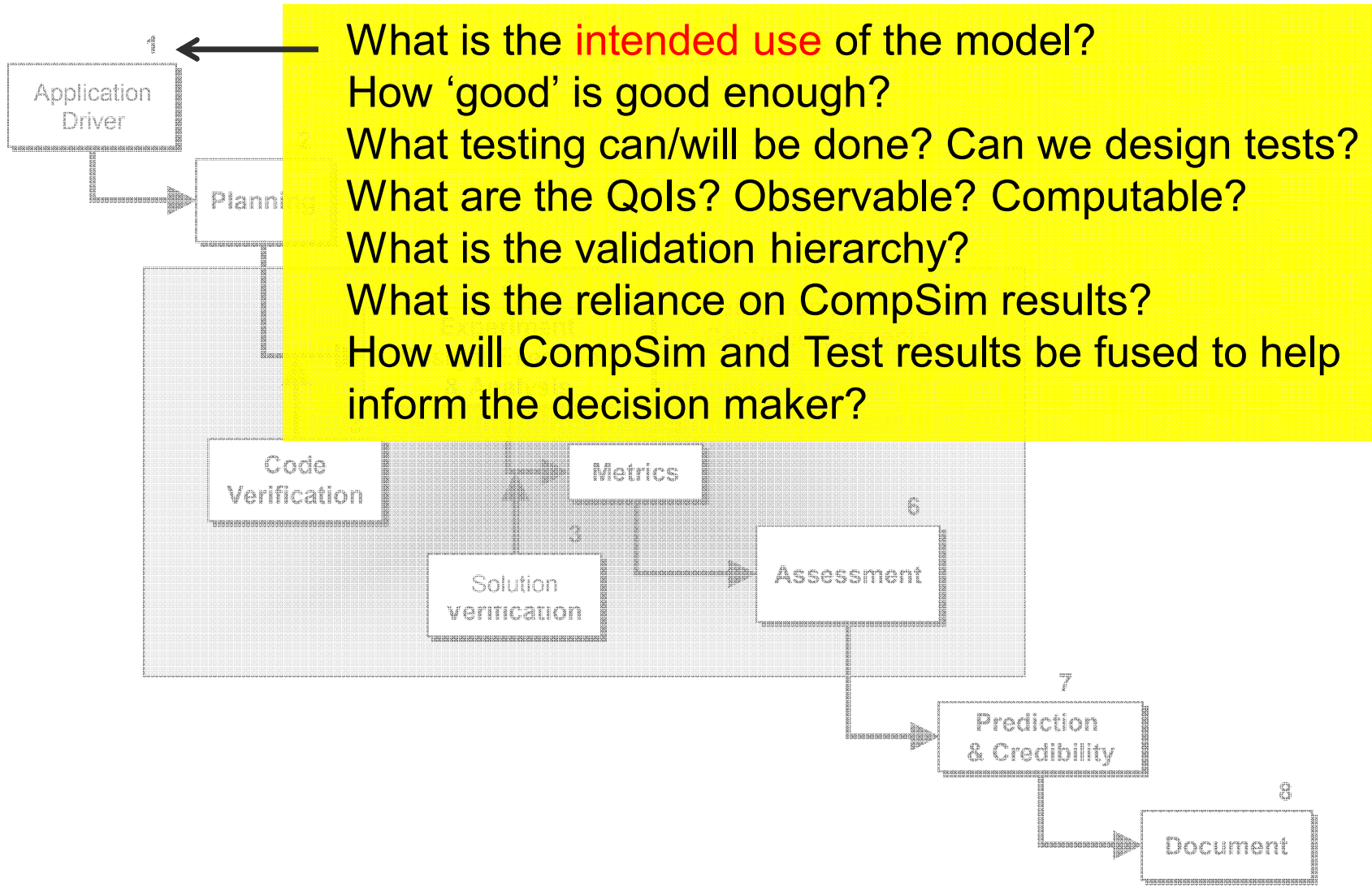
Key Issues:

- Some analysts do some of these activities formally/informally inconsistently
- Amount of formal V&V needed is driven by customer needs
- Internal and external drivers

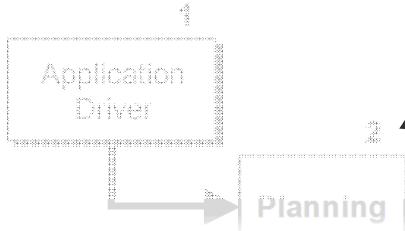


Goal: Assess the credibility of a model prediction for a particular application (a specific intended use)

Understand the Application and Requirements



Assess Capabilities, Identify Gaps & Prioritize Work



Utilize Phenomena Identification and Ranking Tables (PIRTs) and Prediction Capability Maturity Model (PCMM) Tools



PIRT for Example System in Example Scenario

	Importance	Adequacy		
		Math Model	Code	Validation
Aerodynamic/Acoustic Phenomena				
Phenomena 1	M	H	M	L
Phenomena 2				
Phenomena 3	H	M	L	L
Phenomena 4	M	M	M	
Phenomena 5	M	M	L	
Phenomena 6	H	M	L	
Phenomena 7	M	M	L	
Phenomena 8	H	M	M	
Phenomena 9	H	H	M	
Phenomena 10	M	M	M	
Phenomena 11	H	H	H	
Phenomena 12	M	L	L	
Phenomena 13	M	M	M	
Phenomena 14	M	H	H	
Phenomena 15	M	L	L	
Structural Dynamic Phenomena				
Phenomena 1	M	M	L	
Phenomena 2	H	M	M	
Phenomena 3	M	M	L	
Phenomena 4	H	H	M	
Other Important Modeling Issues				
Phenomena 1	H	L	L	
Phenomena 2	M	N/A	N/A	
Phenomena 3	H	N/A	N/A	

ELEMENT	MATURITY				
	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 What features are neglected because of simplifications or stylizations?	<ul style="list-style-type: none"> Judgment only Little or no representational or geometric fidelity for the system and BCs 	<ul style="list-style-type: none"> Significant simplification or stylization of the system and BCs Geometry or representation of major components is defined 	<ul style="list-style-type: none"> Limited simplification or stylization of major components and BCs Geometry or representation is well defined for major components and some minor components Some peer review conducted 	<ul style="list-style-type: none"> 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 Independent peer review conducted 	
Physics and Material Model Fidelity How fundamental are the physics and material models and what is the level of model calibration?	<ul style="list-style-type: none"> Judgment only Model forms are either unknown or fully empirical Few, if any, physics-informed models No coupling of models 	<ul style="list-style-type: none"> Some models are physics based and are calibrated using data from related systems Minimal or ad hoc coupling of models 	<ul style="list-style-type: none"> Physics-based models for all important processes Significant calibration needed using separate effects tests (SETs) and integral effects tests (IETs) One-way coupling of models Some peer review conducted 	<ul style="list-style-type: none"> All models are physics based Minimal need for calibration using SETs and IETs Sound physical basis for extrapolation and coupling of models Full, two-way coupling of models Independent peer review conducted 	
Code Verification Are algorithm deficiencies, software errors, and poor SQE practices corrupting the simulation results?	<ul style="list-style-type: none"> Judgment only Minimal testing of any software elements Little or no SQE procedures specified or followed 	<ul style="list-style-type: none"> Code is managed by SQE procedures Unit and regression testing conducted Some comparisons made with benchmarks 	<ul style="list-style-type: none"> Some algorithms are tested to determine the observed order of numerical convergence Some features & capabilities (F&C) are tested with benchmark solutions Some peer review conducted 	<ul style="list-style-type: none"> All important algorithms are tested to determine the observed order of numerical convergence All important F&Cs are tested with rigorous benchmark solutions Independent peer review conducted 	
Solution Verification Are numerical solution errors and human procedural errors corrupting the simulation results?	<ul style="list-style-type: none"> Judgment only Numerical errors have an unknown or large effect on simulation results 	<ul style="list-style-type: none"> Numerical effects on relevant SRQs are qualitatively estimated Input/output (I/O) verified only by the analysts 	<ul style="list-style-type: none"> Numerical effects are quantitatively estimated to be small on some SRQs I/O independently verified Some peer review conducted 	<ul style="list-style-type: none"> Numerical effects are determined to be small on all important SRQs Important simulations are independently reproduced Independent peer review conducted 	
Model Validation How carefully is the accuracy of the simulation and experimental results assessed at various tiers in a validation hierarchy?	<ul style="list-style-type: none"> Judgment only Few, if any, comparisons with measurements from similar systems or applications 	<ul style="list-style-type: none"> Quantitative assessment of accuracy of SRQs not directly relevant to the application of interest Large or unknown experimental uncertainties 	<ul style="list-style-type: none"> Quantitative assessment of predictive accuracy for some key SRQs from IETs and SETs Experimental uncertainties are well characterized for most SETs, but poorly known for IETs Some peer review conducted 	<ul style="list-style-type: none"> Quantitative assessment of predictive accuracy for all important SRQs from IETs and SETs at conditions/geometries directly relevant to the application Experimental uncertainties are well characterized for all IETs and SETs Independent peer review conducted 	
Uncertainty Quantification and Sensitivity Analysis How thoroughly are uncertainties and sensitivities characterized and propagated?	<ul style="list-style-type: none"> Judgment only Only deterministic analyses are conducted Uncertainties and sensitivities are not addressed 	<ul style="list-style-type: none"> Aleatory and epistemic (A&E) uncertainties propagated, but without distinction Informal sensitivity studies conducted Many strong UQ/SA assumptions made 	<ul style="list-style-type: none"> A&E uncertainties segregated, propagated and identified in SRQs Quantitative sensitivity analyses conducted for most parameters Numerical propagation errors are estimated and their effect known Some strong assumptions made Some peer review conducted 	<ul style="list-style-type: none"> A&E uncertainties comprehensively treated and properly interpreted Comprehensive sensitivity analyses conducted for parameters and models Numerical propagation errors are demonstrated to be small No significant UQ/SA assumptions made Independent peer review conducted 	

- The Predictive Capability Maturity Model (PCMM) is a communication tool for informing stakeholders of the level of maturity of an application-specific simulation capability
 - It is a multidimensional, qualitative metric
 - Determine readiness for stockpile issues
 - Identify gaps in credibility of application
 - Measure progress of integrated simulation effort

PCMM Elements

1. RGF: Representation or Geometric Fidelity
Are representation errors corrupting simulation results?
2. PMMF: Physics and Material Model fidelity
How accurate are the physics and material models?
3. CVER: Code Verification (inc. SQE)
Are software errors or algorithm deficiencies corrupting simulations?
4. SVER: Solution Verification
Are human procedural errors or numerical solution errors corrupting simulation results?
5. VAL: Model Validation
How accurate are the integrated physics and material models?
6. UQ: Uncertainty Quantification/Sensitivity Analysis
What is the impact of variabilities and uncertainties on system performance and margins?

PCMM allows to qualitatively measure our CompSim “due diligence”

PCMM is intended to be a communication and a planning tool

It is not intended to be a report card

PCMM Tool

		Classification Level:	OUO/SRD/etc...			
1						
2		Level could be 0 to 3, integer values not required				
3	Model:		Date:			
4	Lead Assessor:					
5	Team:					
6	Application:					
7						
	Element/Subelement	Desired target level	Level achieved	Is achieved level adequate for intended use	Evidence Links	Comments
8						
9	Code Verification (CVER)					
10	CVER1	Apply Software Quality Engineering (SQE) processes				
11	CVER2	Provide test coverage information				
12	CVER3	Identification of code or algorithm attributes, deficiencies and errors				
13	CVER4	Verify compliance to Software Quality Engineering (SQE) processes				
14	CVER5	Technical review of code verification activities				
15						
16						
17	Physics and Material Model Fidelity (PMMF)					
18	PMMF1	Characterize completeness versus the PIRT				
19	PMMF2	Quantify model accuracy (i.e., separate effects model validation)				
20	PMMF3	Assess interpolation vs. extrapolation of physics and material model				
21	PMMF4	Technical review of physics and material models				
22						
23						
24	Representation and Geometric Fidelity (RGF)					
25	RGF1	Characterize Representation and Geometric Fidelity				
26	RGF2	Geometry sensitivity				
27	RGF3	Technical review of representation and geometric fidelity				
28						
29						
30	Solution Verification (SVER)					
31	SVER1	Quantify numerical solution errors				
32	SVER2	Quantify Uncertainty in Computational (or Numerical) Error				
33	SVER3	Verify simulation input decks				
34	SVER4	Verify simulation post-processor inputs decks				
35	SVER5	Technical review of solution verification				
36						
37						
38	Validation (VAL)					
39	VAL1	Define a validation hierarchy				
40	VAL2	Apply a validation hierarchy				
41	VAL3	Quantify physical accuracy				
42	VAL4	Validation domain vs. application domain				
43	VAL5	Technical review of validation				
44						
45						
46	Uncertainty Quantification (UQ)					
47	UQ1	Aleatory and epistemic uncertainties identified and characterized.				
48	UQ2	Perform sensitivity analysis				
49	UQ3	Quantify impact of uncertainties from UQ1 on quantities of interest				
50	UQ4	UQ aggregation and roll-up				
51	UQ5	Technical review of uncertainty quantification				
52						

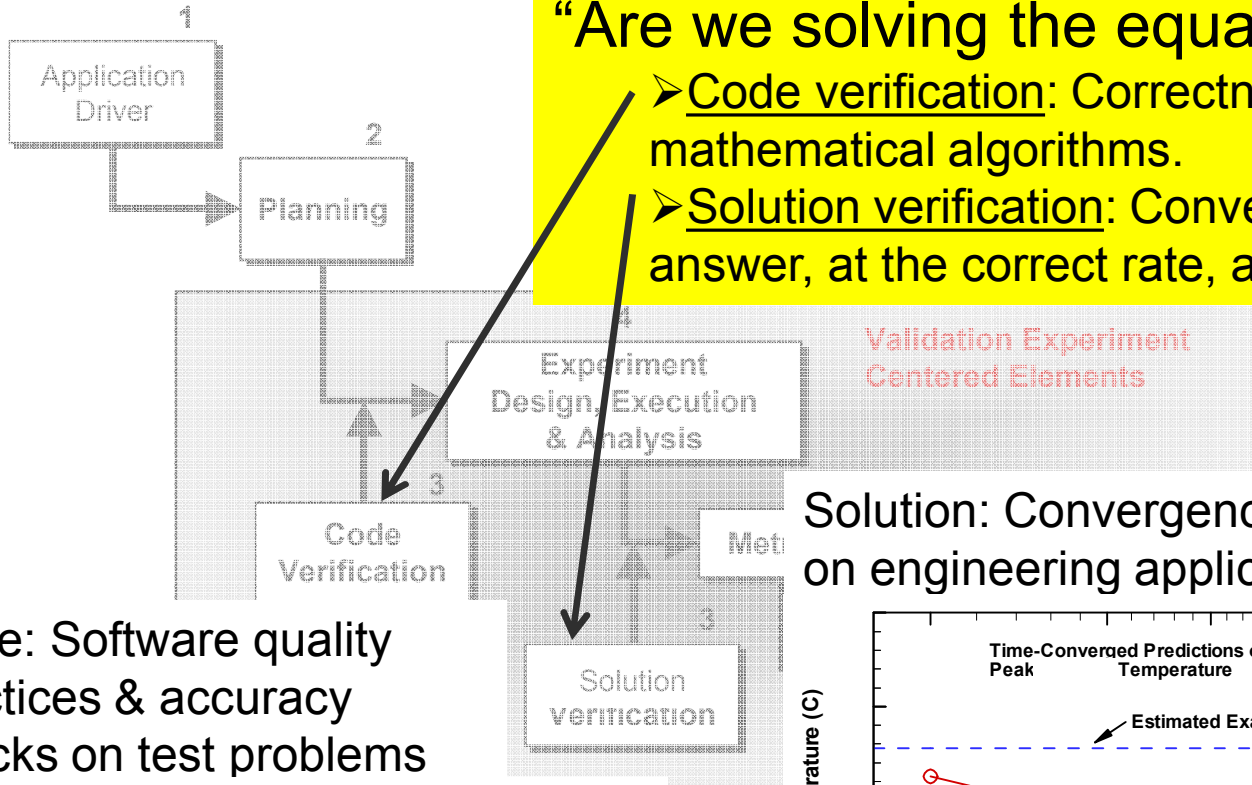
Code Verification (CVER)		Descriptor	
Level 0	No identified SQE process		This element requires
Level 1	Code capability is managed to identified SQE practices		
Level 2	Code capability is managed to identified SQE practices. SQE process is managed.		
Level 3	Code capability is managed to identified SQE practices. SQE process is managed and optimized.		
CVER2: Provide test coverage information		Descriptor	
Level 0	No test coverage reported		
Level 1	Regression testing and/or limited verification tests (VERTS) reported		VERTS = Verification of Software Quality Engineering testing techniques. Answers after too much time spent on a task. Virtually nothing to report. In this sense, a definition of VERTS is required. Here we require software engineers to report VERTS testing results to physics/engineering.
Level 2	Regression testing and VERTS testing with VERTS test feature coverage		

Assessment team should include analysts, code developers, experimentalists, customer and a V&V partner

Verification Activities

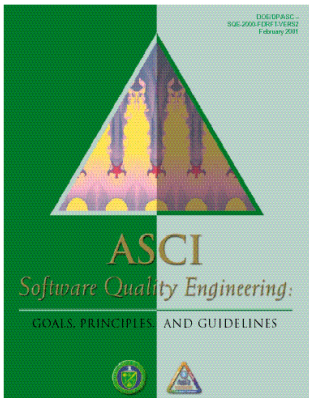
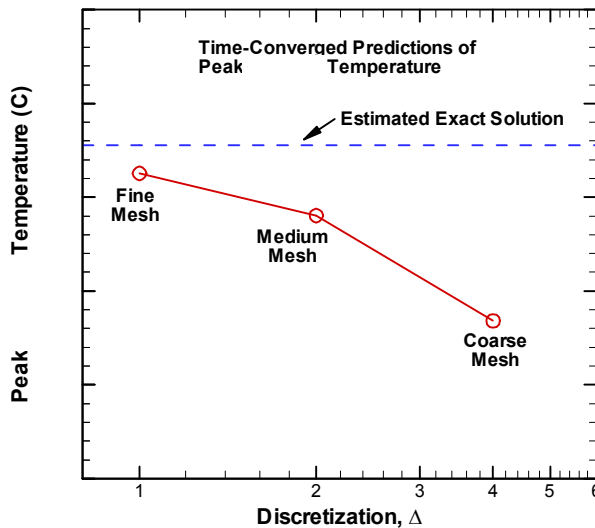
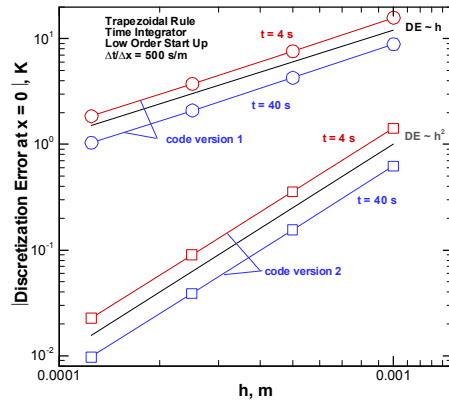
“Are we solving the equations correctly?”

- Code verification: Correctness of implemented mathematical algorithms.
- Solution verification: Convergence to the correct answer, at the correct rate, as model is refined.



Code: Software quality practices & accuracy checks on test problems

Solution: Convergence checks on engineering application



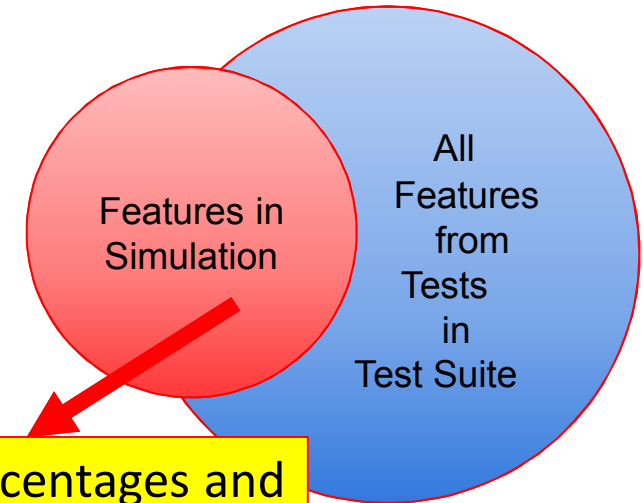
Feature Coverage Tool (FCT) and Sierra

Intersects the set of features used in a simulation with the set from the full suite of nightly and verification tests.

Analyst's simulation (input file)

Annotated by the FCT

```
begin sierra myJob 1214 +
# Define gmres linear system solver, this should not need to be changed.
begin trilinos equation solver gmres 113 +
  param-int AZ_kspace value 1000 4 +
  param-real AZ_ilut_fill value 2 55 + # changed from 5.0
  maximum iterations = 1000 55 +
  solution method = gmres 29 +
  preconditioning method = dd-ilut 7 +
  #matrix scaling = row-sum # added
  preconditioner subdomain overlap = 2 67 +
  residual norm tolerance = 1.0e-9 56 +
end
begin trilinos equation solver direct_solver 113 +
  solution method = amesos-umfpack 10 +
end
end # Define the mesh and which mesh blocks correspond to which elements.
Begin Finite Element Model rect 1214 +
  database Name = single_stack_3D.g 1213 +
  coordinate system is cartesian 79 +
  decomposition method = rcb 173 +
  omit block block_4 8 +
  use material cathode_collector for surface_2 33 +
  use material cathode for block_3 33 +
  use material separator for block_2 33 +
  use material anode for block_1 33 +
End Finite Element Model rect
Begin Output Scheduler forResults 18 +
  At step 0, Increment = 1 16 +
```



Report percentages and annotate the simulation input on intersection

verified

- * one-way: 58%
- * two-way: 27%

tested

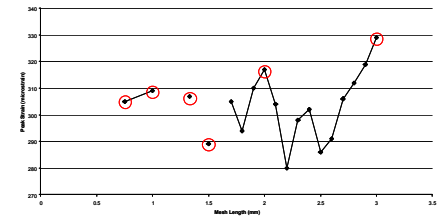
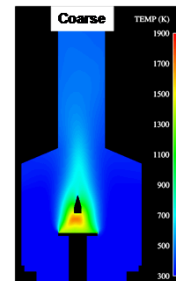
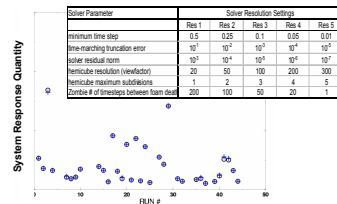
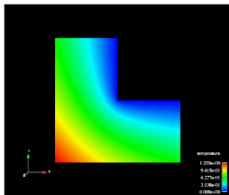
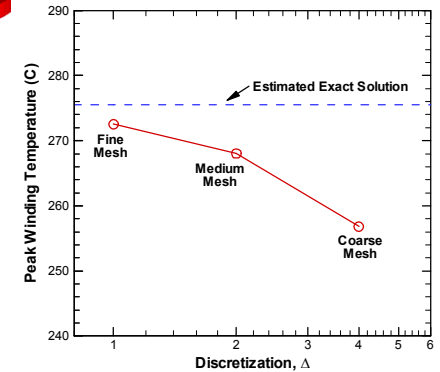
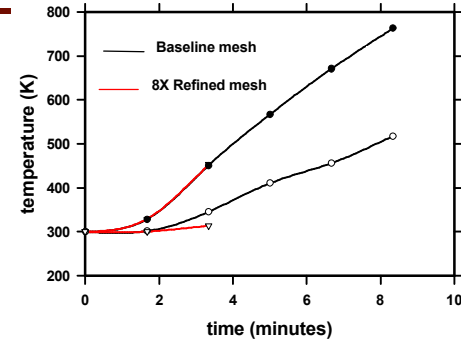
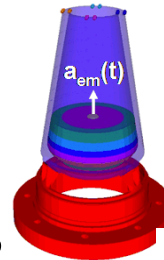
- * one-way: 71%

untested

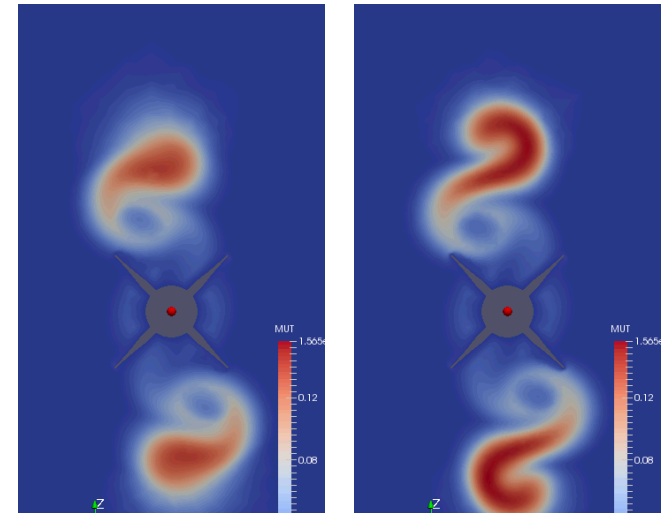
- * ignored

Solution Verification

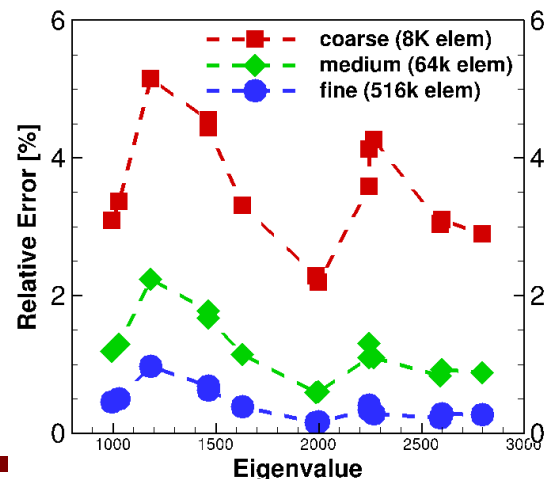
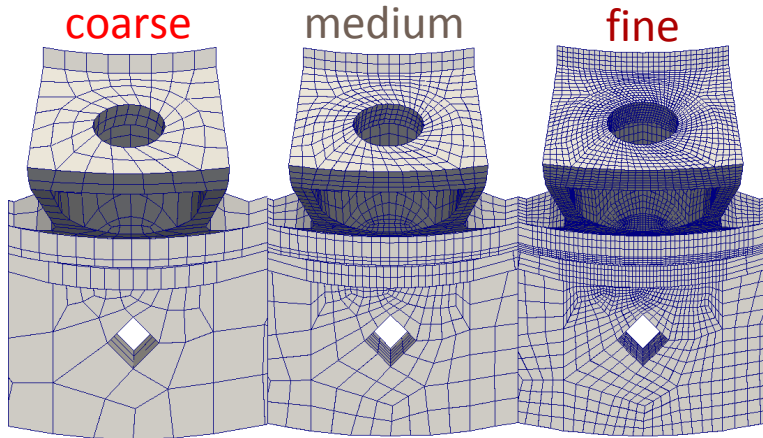
- Quantify numerical solution errors
 - What is the impact of numerical solution errors on relevant system response quantities or QoIs?
- Verify all simulation inputs and outputs
 - Have we corrupted simulation results with incorrect inputs or post processing errors?



- Provide a scalable software package to help automate convergence analysis.
- Create uniformly and adaptively refined sequences of unstructured grids.



Rear view of vortices arising from jet in cross flow using coarse mesh (left) and offline adapted mesh (3x).



- Errors from extrapolation (first 14 nonzero frequencies)
- This allows the assessment of the accuracy of each mesh

V&V/UQ tools and Credibility (via PCMM)



Compare

Representation and Geometric Fidelity (RGF)

RGF1	Characterize Representation and Geometric Fidelity	2	1
RGF2	Geometry sensitivity	3	2
RGF3	Technical review of representation and geometric fidelity	2	1

FCT

Code Verification (CVER)

CVER1	Apply Software Quality Engineering (SQE) processes	2	2
CVER2	Provide test coverage information	2	2
CVER3	Identification of code or algorithm attributes, deficiencies and errors	2	1
CVER4	Verify compliance to Software Quality Engineering (SQE) processes	2	2
CVER5	Technical review of code verification activities	2	2

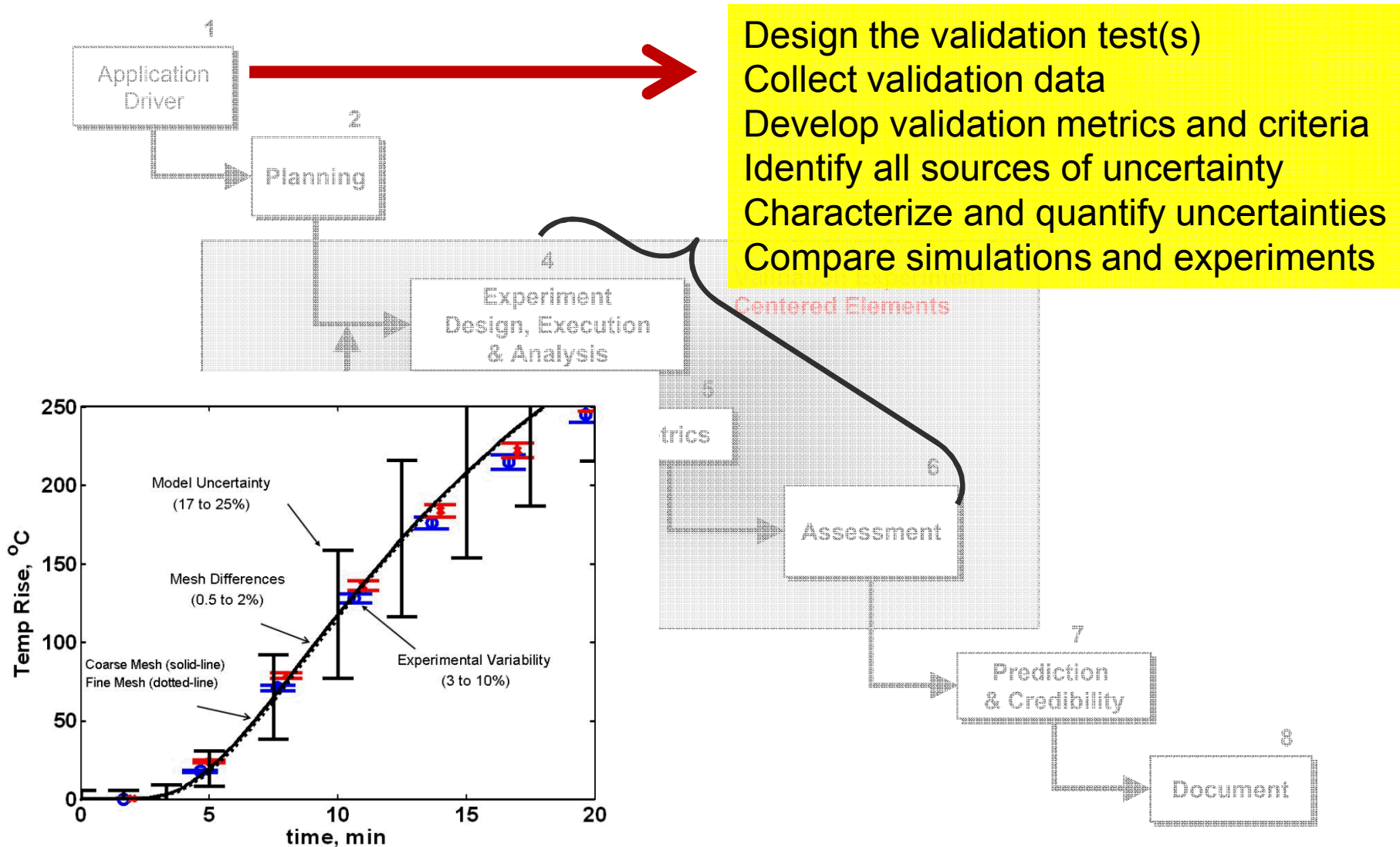
Percept

Solution Verification (SVER)

SVER1	Quantify numerical solution errors	2	1
SVER2	Quantify Uncertainty in Computational (or Numerical) Error	3	2
SVER3	Verify simulation input decks	3	3
SVER4	Verify simulation post-processor inputs decks	2	1
SVER5	Technical review of solution verification	3	2

These tools will help a user address these items as well as provide evidence.

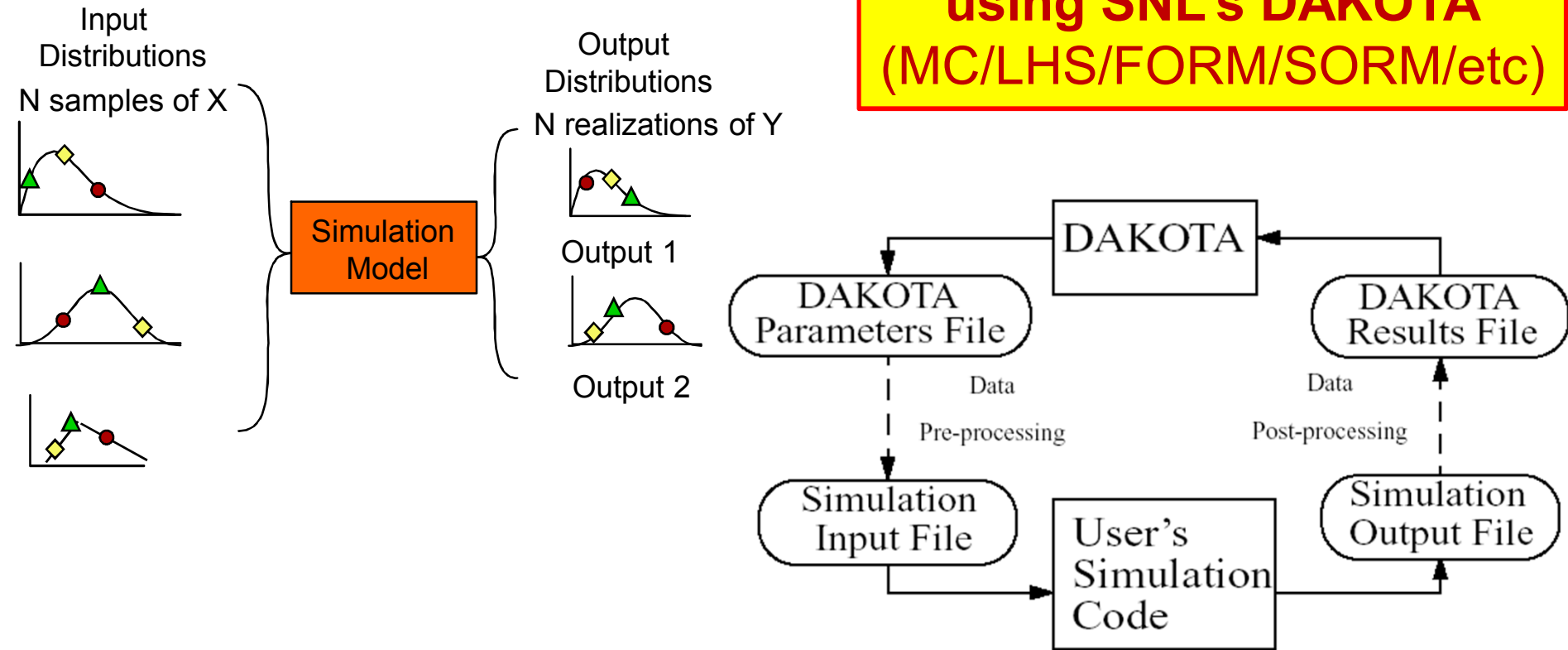
Validation/UQ Activities



Uncertainty Quantification (UQ)

- Statistical propagation of uncertainty through a simulation model, and statistical interpretation of model response.

This process is enabled using SNL's DAKOTA (MC/LHS/FORM/SORM/etc)

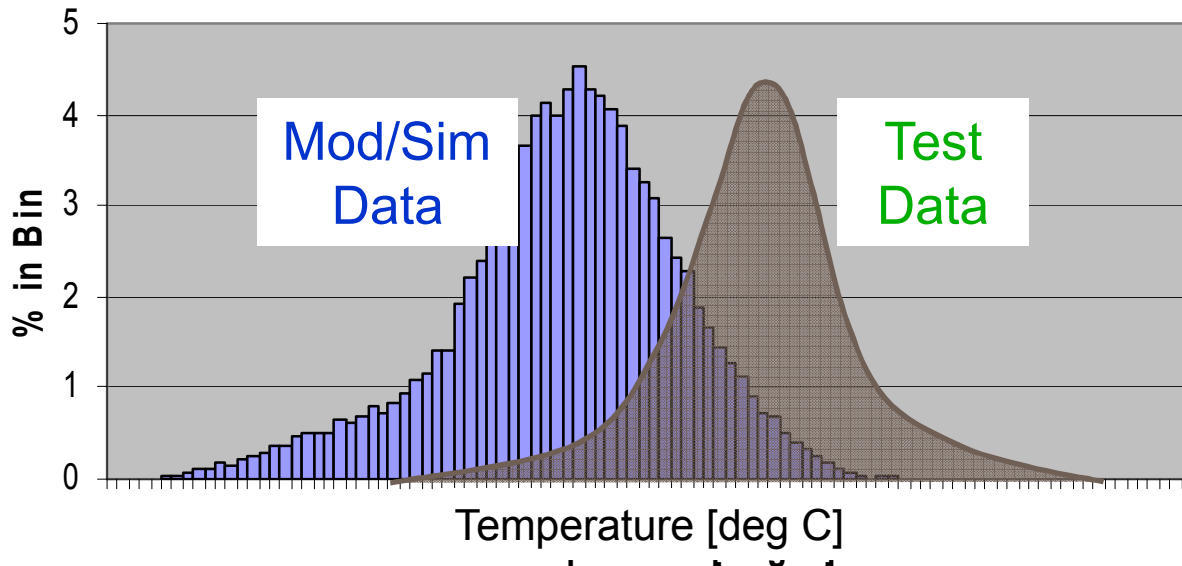


- The **process** of determining the degree to which a model is an **accurate** representation of the **real world** from the perspective of the **intended uses** of the model
- “Are we solving the correct equations?”

Some observations on validation:

- Uncertainty in both model simulations and experimental data should be included
- A criteria for determining the required accuracy needs to be defined *a priori* and should be relative to the intended use of the model
- Comparing 2 lines on a plot (“vugraph norm”) is a way of doing validation but **not** very rigorous

Final Temperature Values



Uncertainty Quantification:

UQ methods generate ensemble of mod/sim data.

Validation:

Compare simulation data to test data w/unc

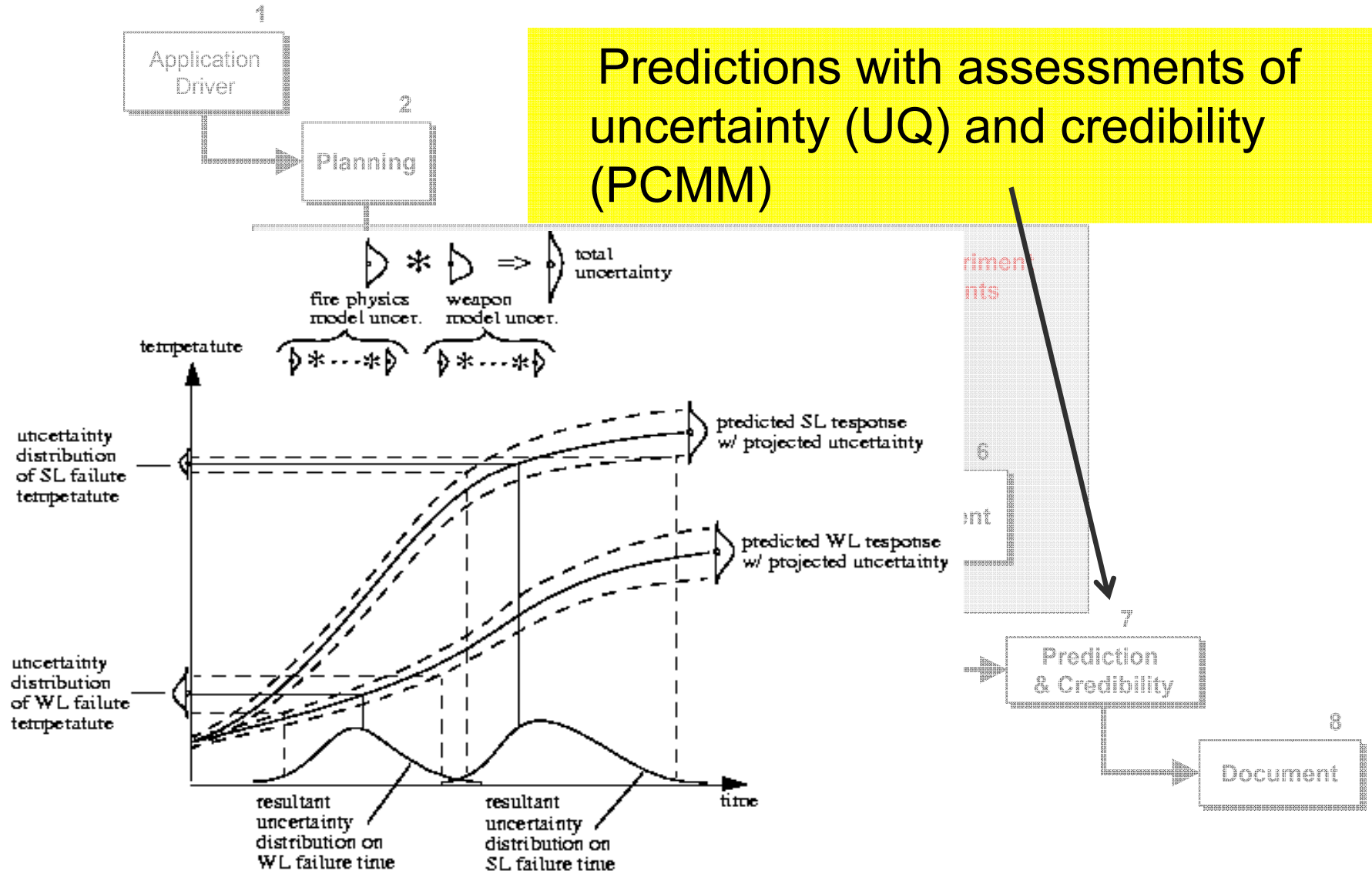
- Quantify amount of “overlap”
- Assess sufficiency of overlap

UQ methods provide statistical info on the code output data:

- Probability distribution on Temperature, given various x_1, \dots, x_N inputs.
- Correlations (trends) of Temperature vs. x_1, \dots, x_N .
- Mean(T), StdDev(T), Probability($T > T_{critical}$)

Predictions

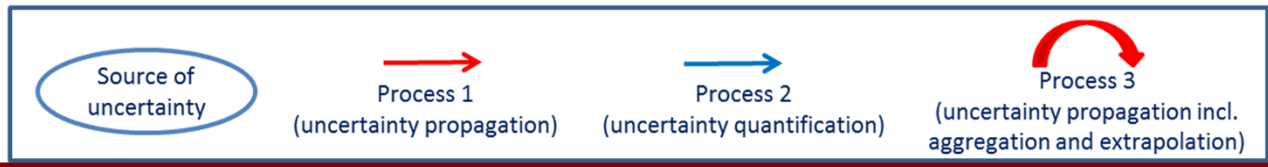
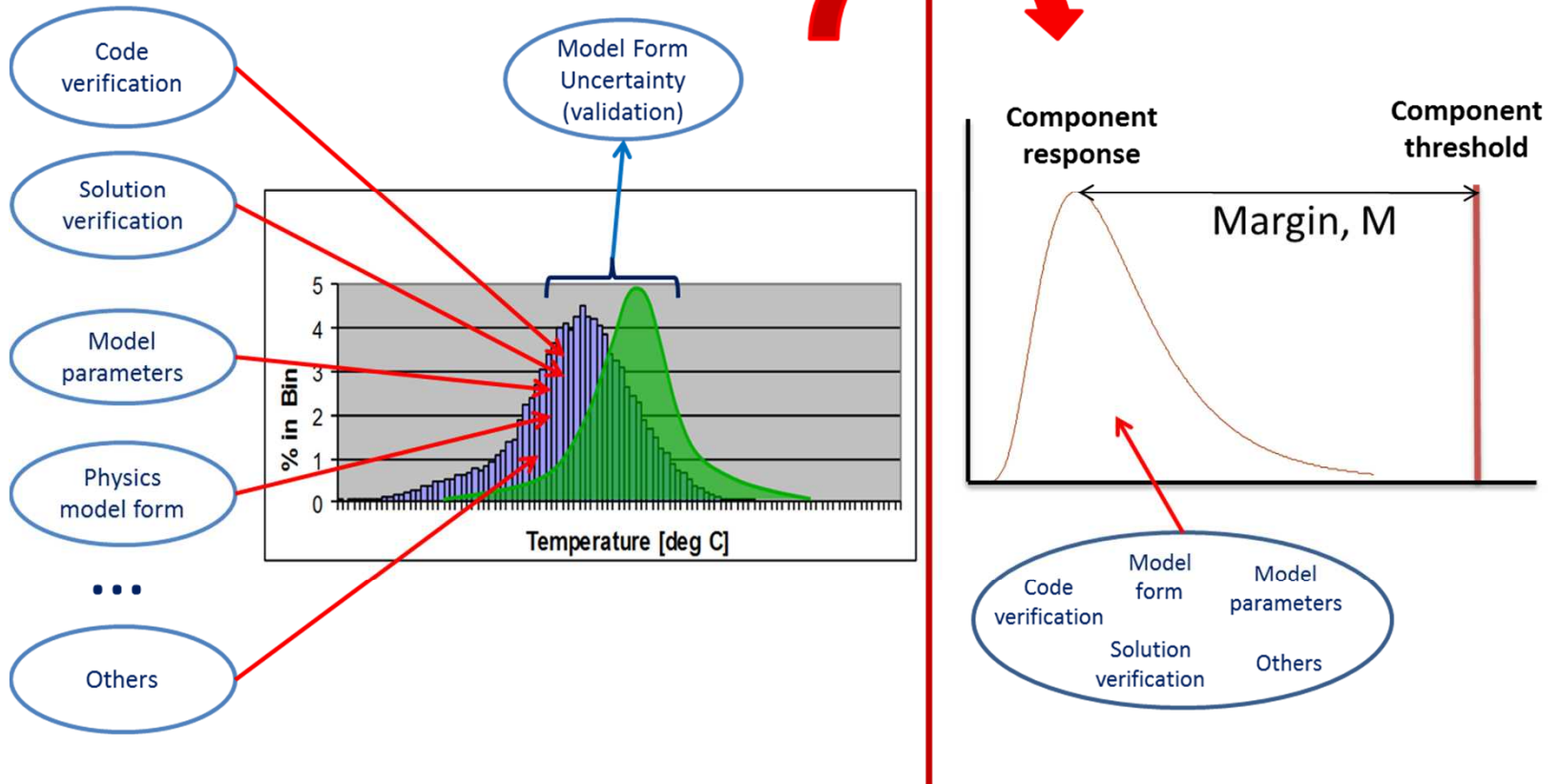
Predictions with assessments of uncertainty (UQ) and credibility (PCMM)



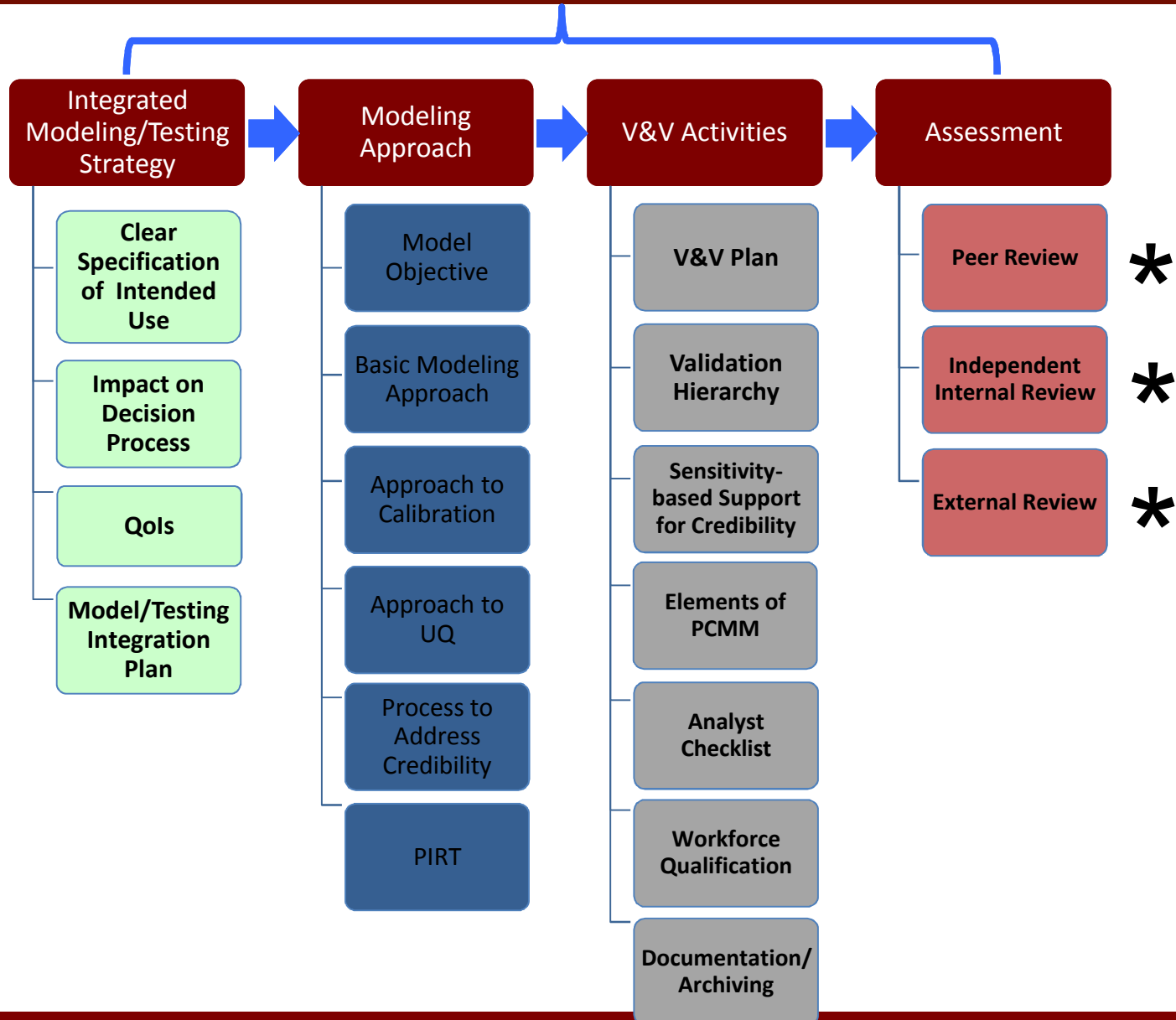
From validation to application space

Validation Space

Application Space
(maybe QMU)



Assessment of Prediction Credibility



- Must budget resources to understand numerical uncertainties
- Need to insert UQ earlier into model building stage
- Need tighter integration between CompSim and testing
- Must develop tractable models to perform UQ studies
- Maximize information gained with limited, precious exptl testing
 - Have a testing program that continuously improves including diagnostics
- SQA/SQE of workflows
- Pedigreed Material Databases are invaluable
- No single point responses – every result should have error bars
- Management assurance vs quality assurance
 - Peer reviews, data reviews, code reviews, model reviews
 - Human factors/choices can dominate uncertainties

- Must understand what is the intended use of CompSim
- Need practical, graded approaches wrt program constraints and expectations wrt achievable CompSim credibility
- Should present CompSim results along with credibility evidence and the associated risks/limitations
- This all holds for experimental credibility also!!
- Easy to motivate → Hard to do
- Supporting V&V/UQ activities must start at the beginning
- Peer reviews are essential
- One of the main reasons for having a V&V/UQ process is to increase the confidence in CompSim results
 - PCMM is a infrastructure to communicate this confidence and its supporting evidence

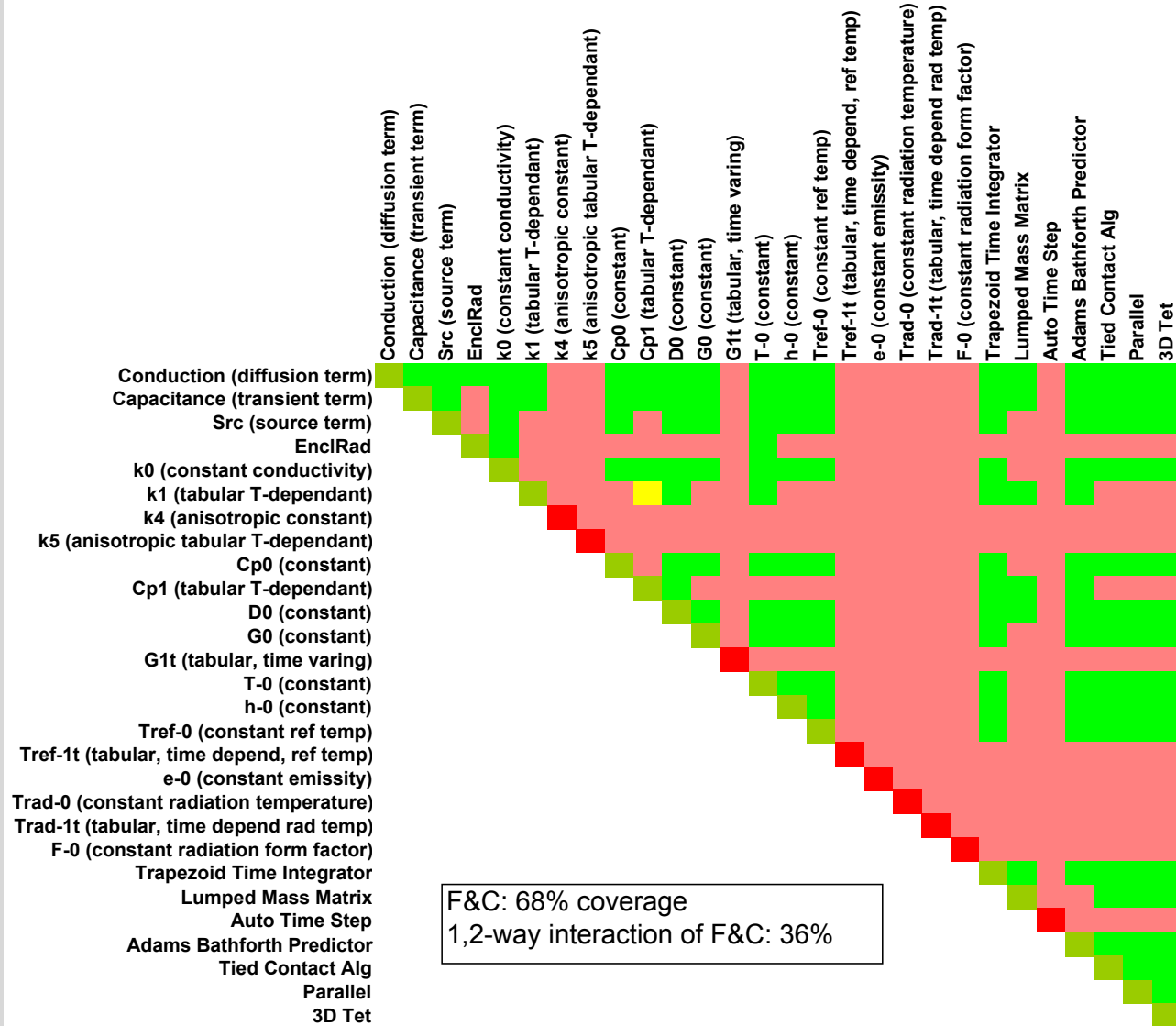
THANK YOU

Extra Slides



Feature Coverage Tool (FCT)

- An efficient way to judge how well features are tested and help ensure proper usage of features
- A metric to communicate gaps to development team
- Evidence for the credibility evidence package



- Verification – “Are we solving the equations correctly?”
 - Code verification: Correctness of implemented mathematical algorithms.
 - Solution verification: Convergence to the correct answer, at the correct rate, as model is refined.
- Validation – “Are we solving the right equations?”
 - Correctness of physical models and sufficiency for the application.
 - Model Validation is the **process** of determining the degree to which a model is an **accurate** representation of the **real world** from the perspective of the **intended uses** of the model
- Uncertainty Quantification (UQ):
 - Statistical propagation of uncertainty through a simulation model, and statistical interpretation of model response.
- Quantification of Margins and Uncertainties (QMU):
 - Using the simulation model to make system performance predictions with quantified uncertainty, and with quantified margins with respect to system performance requirements.

- Program integration
 - Improving communication and integration between S&T with NW communities
 - Improving understanding of each others needs and capabilities and restrictions/limitations/gaps
 - Understand the reliance on CompSim vs testing and their fusion to inform decision making and the understanding of risks
 - Understand the customers important Qols for validation
 - Integrated planning and scheduling
 - Tight partnerships between customers, integration leads, analysts, code developers, experimentalists and V&V partner
 - Early negotiation of intended use of CompSim model and results (same for expts)
 - Need practical, graded approaches wrt program constraints and CompSim predictivity
- Systems must understand ramifications of late program changes
 - Model building challenges
 - Re-establishing credibility evidence
 - Reduction of testing information greatly impacts validation