SAND2013-2967P

Intrinsic Verification & Validation/UQ:
The current workflow and the future concept
- Illustrated through an example

W. Witkowski
Validation and Uncertainty Quantification Dept.
Sandia National Laboratories

Verification, Validation, and Uncertainty Quantification are the Science Behind QMU

- Verification – “Are we solving the equations correctly?”
  - Correctness of implemented mathematical algorithms.
  - 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.

Intrinsic V&V/UQ Conceptual Strategy

Intrinsic = "being part of the fundamental nature or substance of something"

Vision

Support the customer who uses CompSim to make *confident* and *reliable* decisions via an inseparable integration of V&V/UQ/QMU analyses into the workflow

- This requires understanding our customer's needs and the customer understanding what he needs

What is different philosophically?

- The workflow determines the appropriate level of rigor with “self-checks”
- Robust to human usage (from novice to seasoned analysts)
  - Analysts rely on intuition from past successes which could be dangerous
  - Currently, there aren’t any good ways to assess all uncertainty/errors in our codes, therefore, the analysts must be protected
- Enabling analysis codes with V&V/UQ capabilities for internal uses
  - Embedded sensitivity analysis
  - Inserting uncertainties straight into application input files
- Process workflow “learns” from previous application experiences
- Process workflow generates supporting credibility evidence
- Take home messages:
  - We must continue to educate our customers on the importance and necessity of V&V/UQ analysis at the appropriate level of rigor being performed along-side of all CompSIm activities.
  - We must invest in capabilities to ensure credible predictions

Current Mission
First steps to Intrinsic

Enable and facilitate the analyst's ability to perform (and *automatically* perform where appropriate) the components of V&V/UQ analysis to insure that the appropriate rigor and consistent credibility is assured and evaluated for high impact numerical predictions with an assessment of the associated margins.

Through:

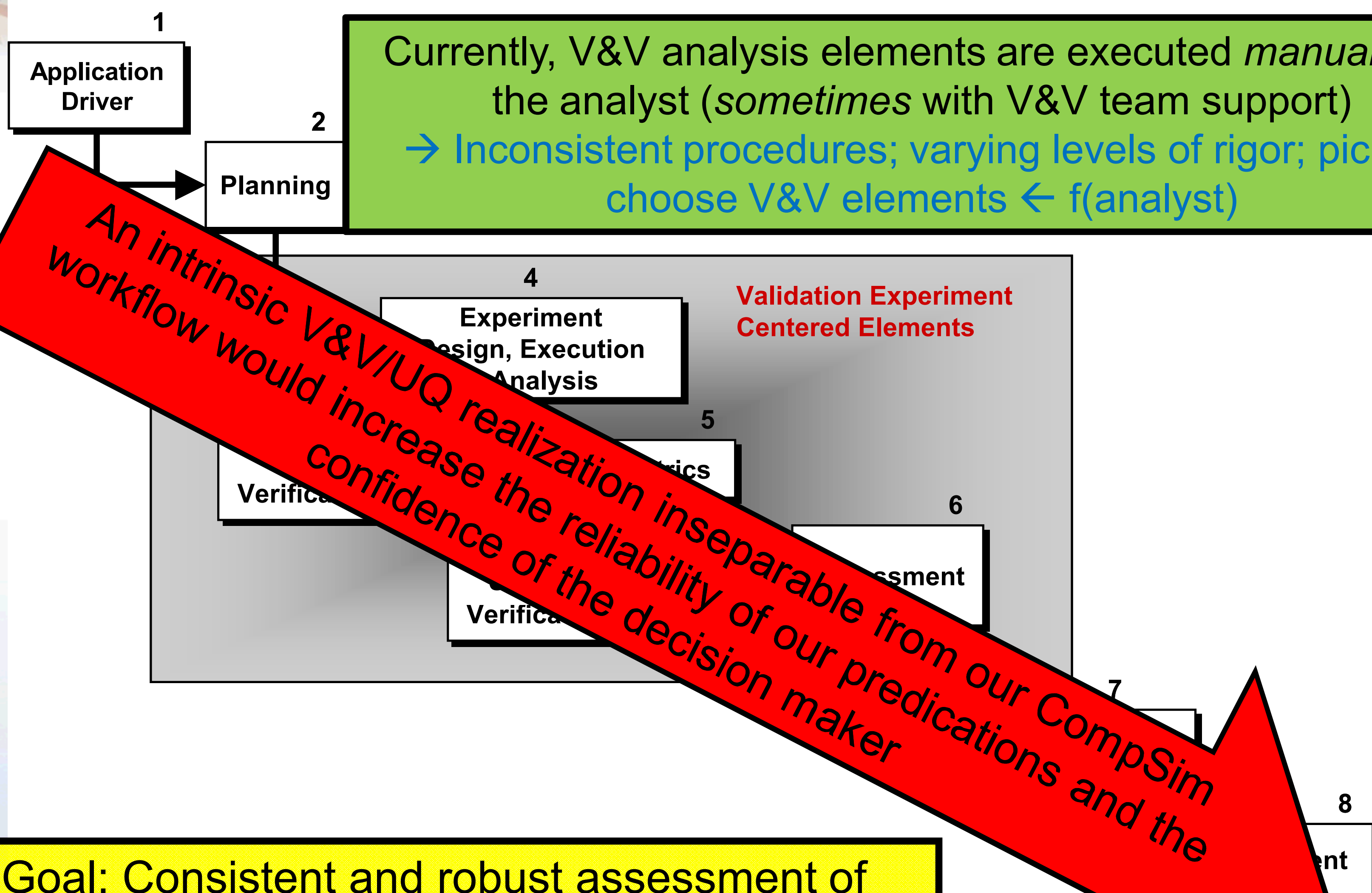
- State-of-the-art methods for verification, validation and uncertainty quantification integrated into the analyst's workflow
- Expert-informed guidance for streamlined V&V/UQ processes
- Data-guided tools to support credible CompSim-based analyses and decision making

Intrinsic concept is complex and difficult to define and relate → illustrate some characteristics and an initial path forward through an example

Sandia National Laboratories

Current Recommended V&V Process Elements

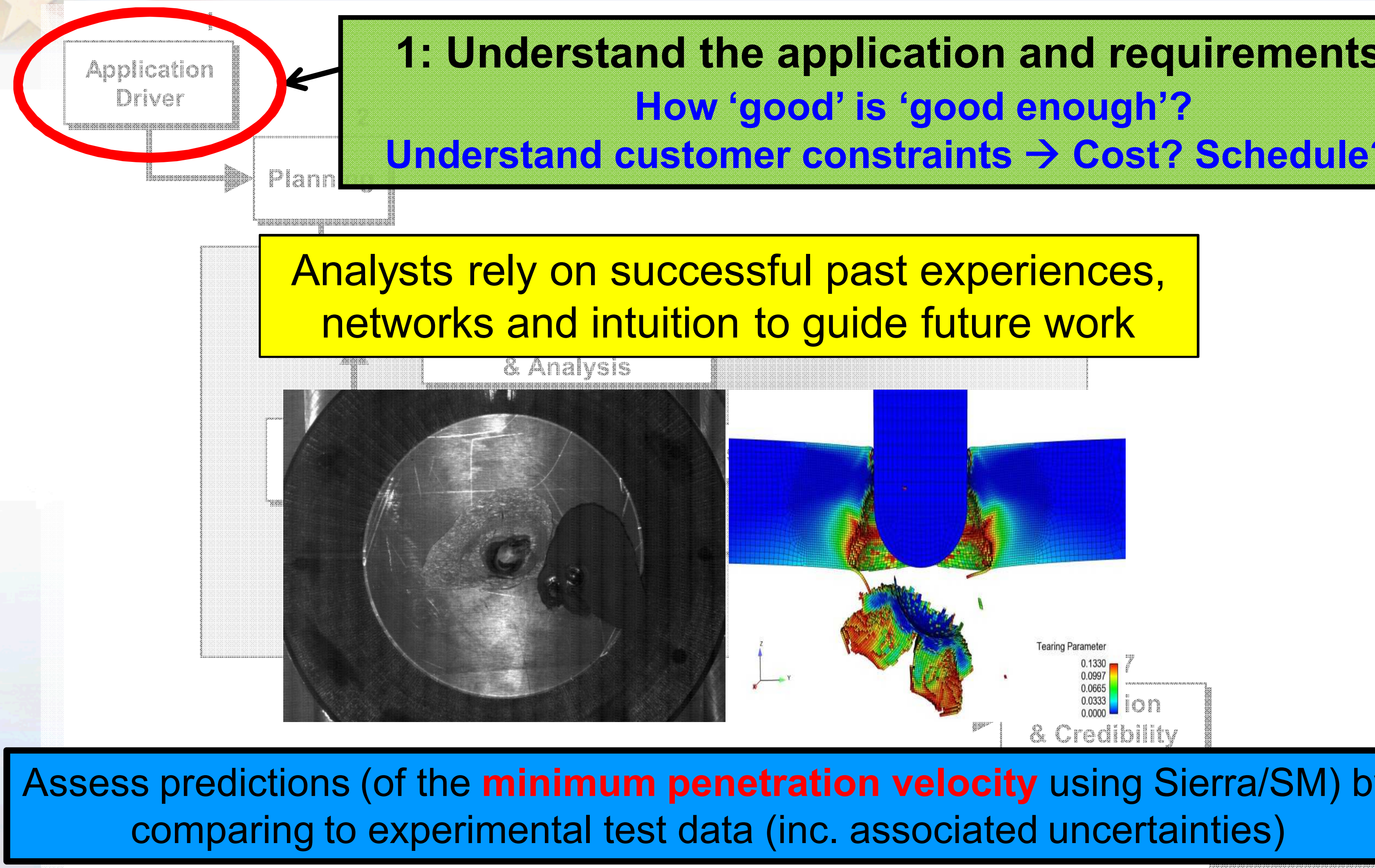
Currently, V&V analysis elements are executed *manually* by the analyst (*sometimes* with V&V team support)
→ Inconsistent procedures; varying levels of rigor; pick-n-choose V&V elements ← f(analyst)



An intrinsic V&V/UQ realization inseparable from our CompSim workflow would increase the reliability of our predications and the confidence of the decision maker

Goal: Consistent and robust assessment of credibility and uncertainty assessment of model predictions for a particular application

Current V&V Process: Understand Application



1: Understand the application and requirements

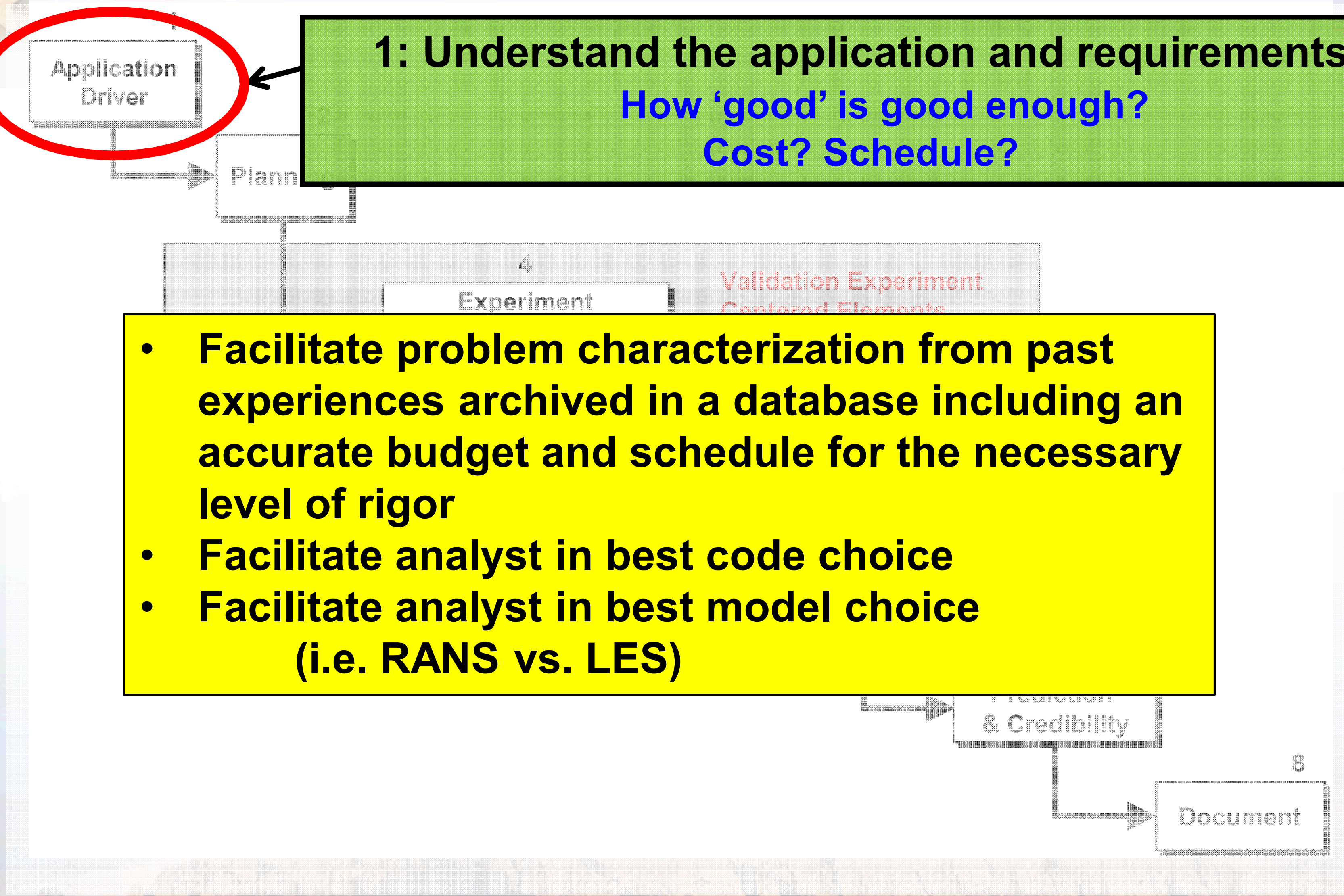
How 'good' is 'good enough'?

Understand customer constraints → Cost? Schedule?

Analysts rely on successful past experiences, networks and intuition to guide future work

Assess predictions (of the **minimum penetration velocity** using Sierra/SM) by comparing to experimental test data (inc. associated uncertainties)

Intrinsic V&V Process: Understand Application



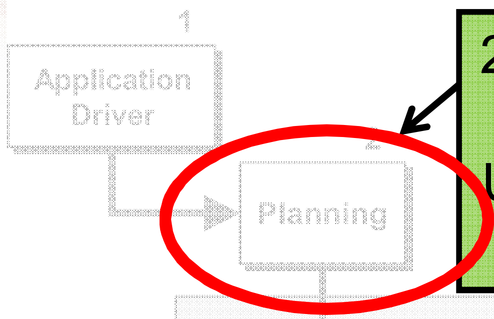
1: Understand the application and requirements

How 'good' is good enough?
Cost? Schedule?

- **Facilitate problem characterization from past experiences archived in a database including an accurate budget and schedule for the necessary level of rigor**
- **Facilitate analyst in best code choice**
- **Facilitate analyst in best model choice (i.e. RANS vs. LES)**

Sandia National Laboratories

Current V&V Process: Planning



2: Assess capabilities, identify gaps, & prioritize work

Utilize Phenomena Identification and Ranking Tables (PIRTs) and PCMM

Phenomena	Consensus	Adequacy		
	Importance	Math Model	Sierra/SM Code	Validation
Large elastic-plastic deformation of metals	H	H	M	M
Ductile material failure	H	M	M	L
Contact	H	H	M	M
Friction between punch and test item	M	M	M	L
Enforcement of boundary conditions	L	H	H	L
Inertial loads	H	H	H	M

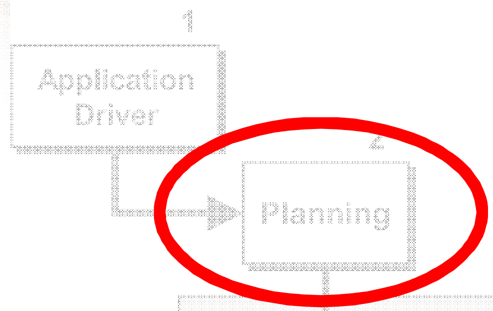
PIRT – evaluated by analysts from their knowledge base

PCMM – incomplete; difficult to understand "maturity levels"; inconsistent evals; beginning to have team evaluations; f(team)

ELEMENT				
Represen[illegible] Geometr[illegible] What features [illegible] because of si[illegible] styliz[illegible]				
Physics a[illegible] Model Fidelity How fundamental are the physics and material models and what is the level of model calibration?	[illegible] unknown or fully empirical • Few, if any, physics-informed models • No coupling of models	[illegible] calibrated using data from related systems • Minimal or ad hoc coupling of models	• Significant calibration needed using separate effects tests (SETs) and integral effects tests (IETs) • One-way coupling of models • Some peer review conducted	[illegible] 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?	• Judgment only • Minimal testing of any software elements • Little or no SQE procedures specified or followed	• Code is managed by SQE procedures • Unit and regression testing conducted • Some comparisons made with benchmarks	• 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	• 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?	• Judgment only • Numerical errors have an unknown or large effect on simulation results	• Numerical effects on relevant SRQs are qualitatively estimated • Input/output (I/O) verified only by the analysts	• Numerical effects are quantitatively estimated to be small on some SRQs • I/O independently verified • Some peer review conducted	• 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?	• Judgment only • Few, if any, comparisons with measurements from similar systems or applications	• Quantitative assessment of accuracy of SRQs not directly relevant to the application of interest • Large or unknown experimental uncertainties	• 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	• 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?	• Judgment only • Only deterministic analyses are conducted • Uncertainties and sensitivities are not addressed	• Aleatory and epistemic (A&E) uncertainties propagated, but without distinction • Informal sensitivity studies conducted • Many strong UQ/SA assumptions made	• 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	• 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

Sandia National Laboratories

Intrinsic V&V Process: Planning



- Updated PCMM and defined process guidelines for more accurate, consistent and useful assessments through Excel-based tool

→ Web-based question-guided tool to facilitate evaluation linked to archived past applications to supply level evaluations

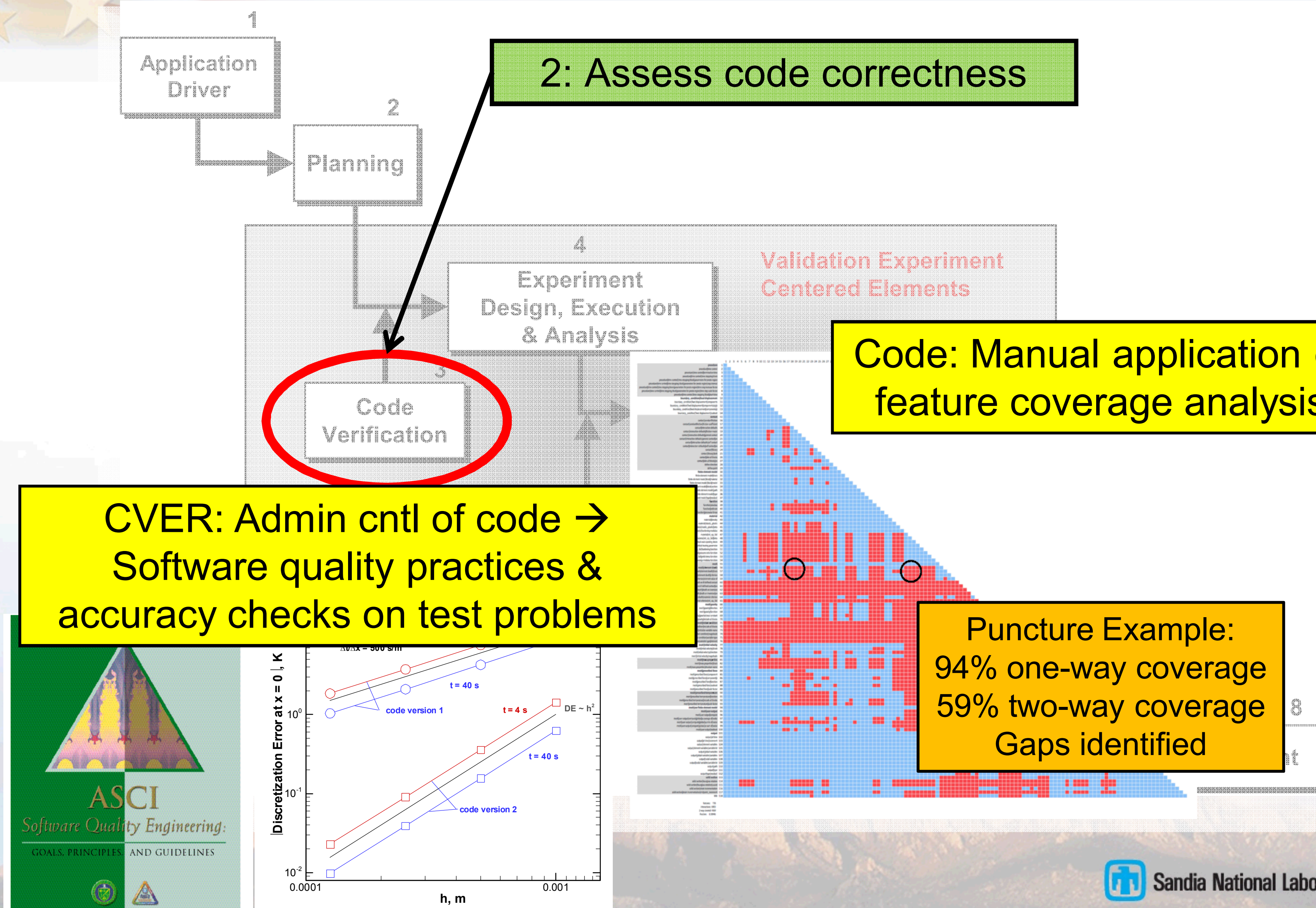
Phenomena	Consensus Importance	Adequacy: Math Model	Sierra/SM Code	
Large elastic-plastic deformation of metals	H	H	M	
Ductile material failure	H	M	M	L
Contact	H	H	M	M
Friction between punch and test item	M	M	M	L
Enforcement of boundary conditions	L	H	H	L
Inertial loads	H	H	H	M

→Extract evidence from previous similar applications and provide adequacy levels automatically

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?	• 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 • Some peer review conducted	• 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?	• Judgment only • Model forms are either unknown or fully empirical • Few if any physics-informed models • No coupling of models	• Some models are physics based and are calibrated using data from related systems • Minimal or ad hoc coupling of models	• 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	• 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?	• Judgment only • Minimal testing of any software elements • Little or no SQE procedures specified or followed	• Code is managed by SQE procedures • Unit and regression testing conducted • Some comparisons made with benchmarks	• 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	• 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?	• Judgment only • Numerical errors have an unknown or large effect on simulation results	• Numerical effects on relevant SRQs are qualitatively estimated • Input/output (I/O) verified only by the analysts	• Numerical effects are quantitatively estimated to be small on some SRQs • I/O independently verified • Some peer review conducted	• 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?	• Judgment only • Few, if any comparisons with measurements from similar systems or applications	• Quantitative assessment of accuracy of SRQs not directly relevant to the application of interest • Large or unknown experimental uncertainties	• 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	• 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?	• Judgment only • Only deterministic analyses are conducted • Uncertainties and sensitivities are not addressed	• Aleatory and epistemic (A&E) uncertainties propagated, but without distinction • Informal sensitivity studies conducted • Many strong UQ/SA assumptions made	• 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	• 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

Sandia National Laboratories

Current V&V Process: Code Verification



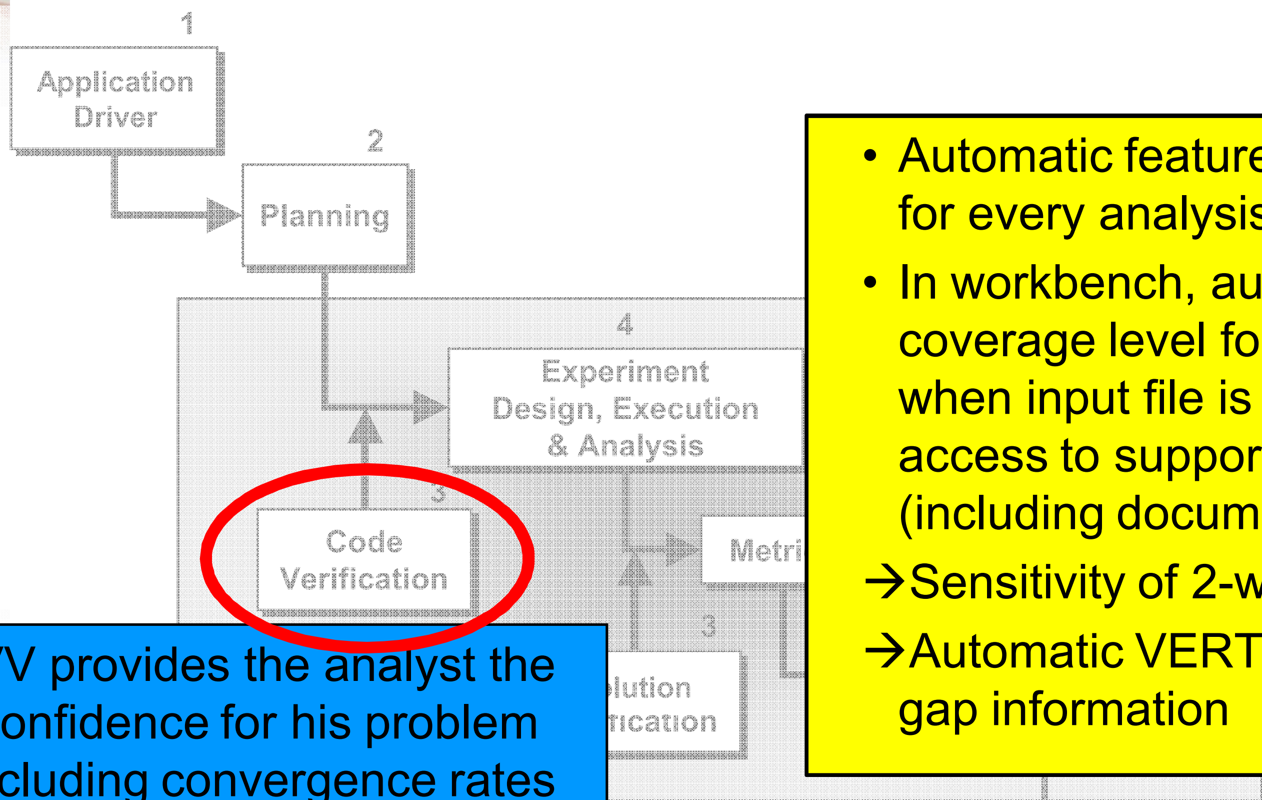
2: Assess code correctness

Code: Manual application of feature coverage analysis

CVER: Admin cntl of code → Software quality practices & accuracy checks on test problems

Puncture Example:
94% one-way coverage
59% two-way coverage
Gaps identified

Intrinsic V&V Process: Code Verification

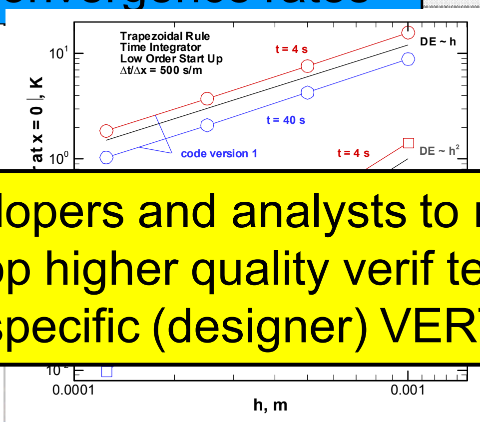


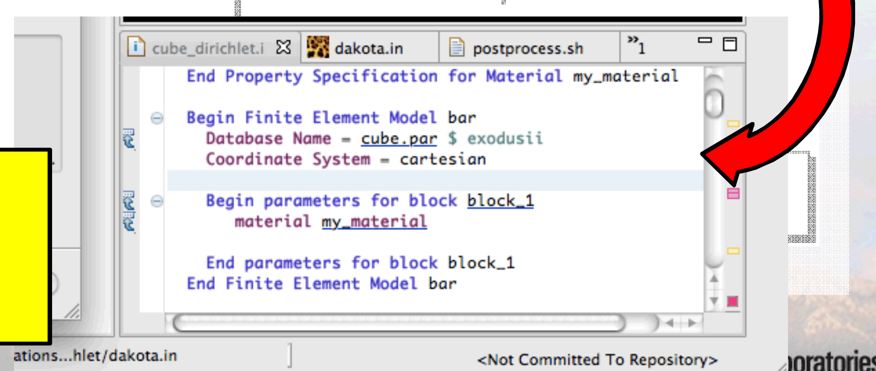
- Automatic feature coverage analysis for every analysis performed.
- In workbench, automatic display of coverage level for each feature when input file is displayed with access to supporting VERTS (including documentation.)

→Sensitivity of 2-way gaps reported

→Automatic VERTS generation from gap information

IVV provides the analyst the confidence for his problem including convergence rates



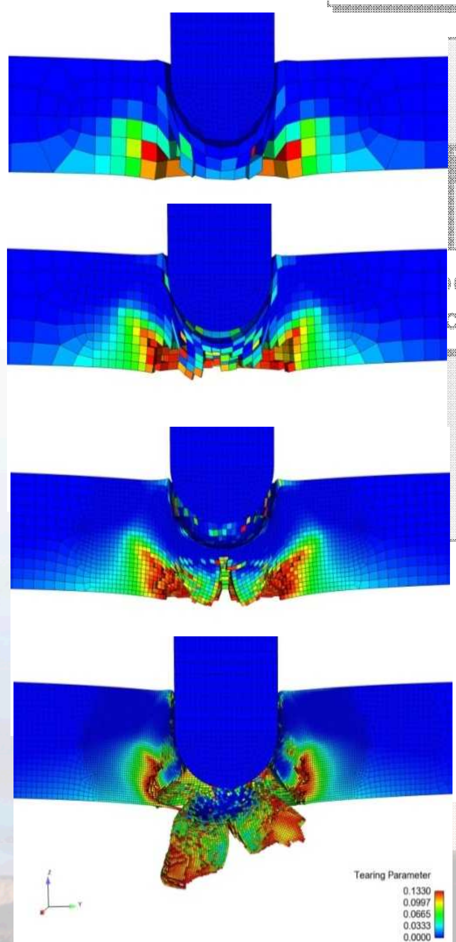


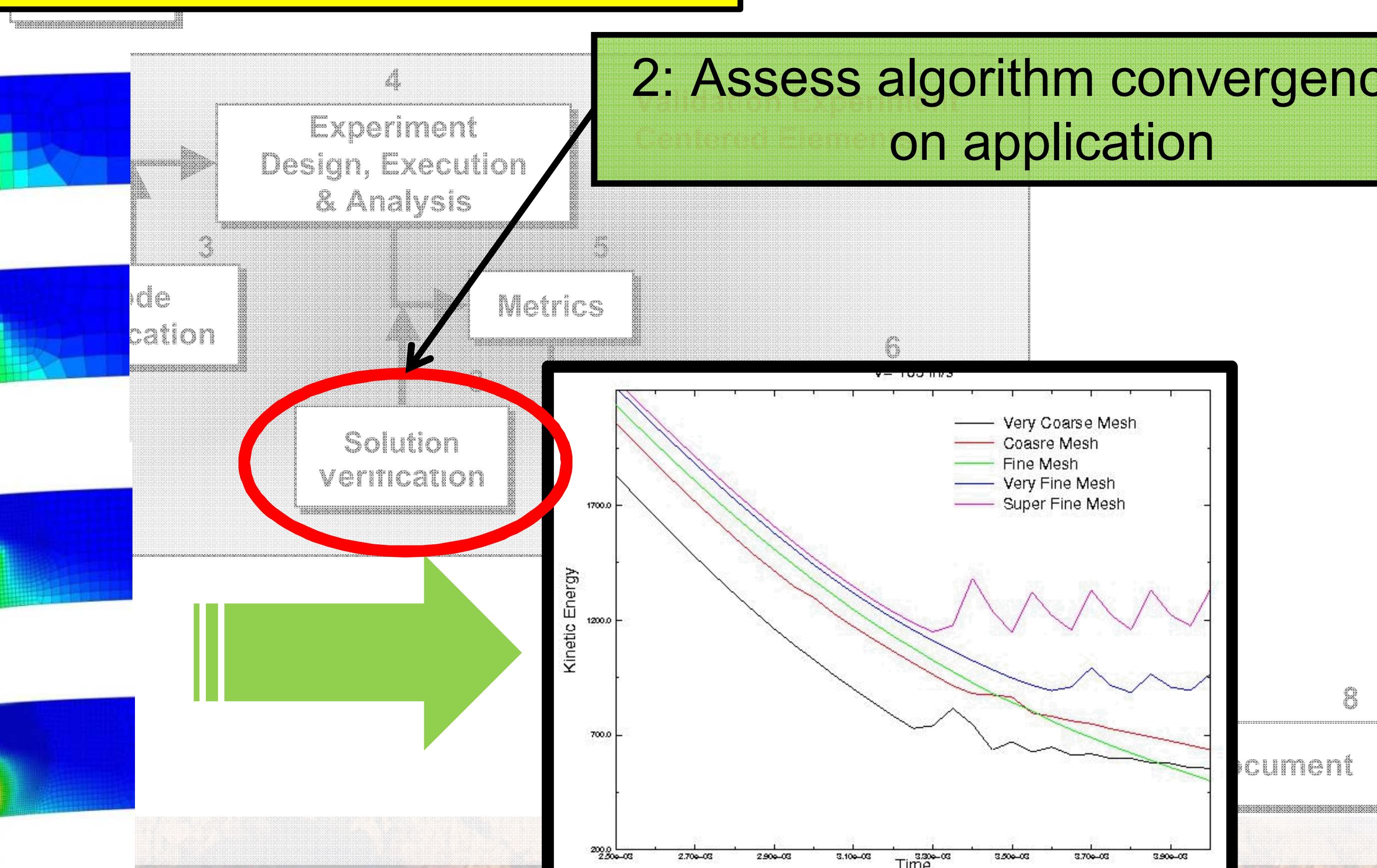
- Enable developers and analysts to more easily develop higher quality verif tests

→ Application-specific (designer) VERTS

Current V&V Process: Solution Verification

- Manual creation of several meshes to perform mesh convergence study
- Manual computation of convergence rates





2: Assess algorithm convergence on application

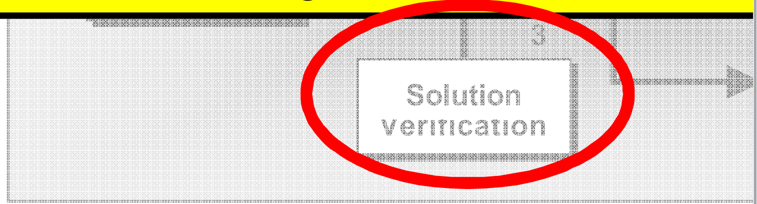
Sandia National Laboratories

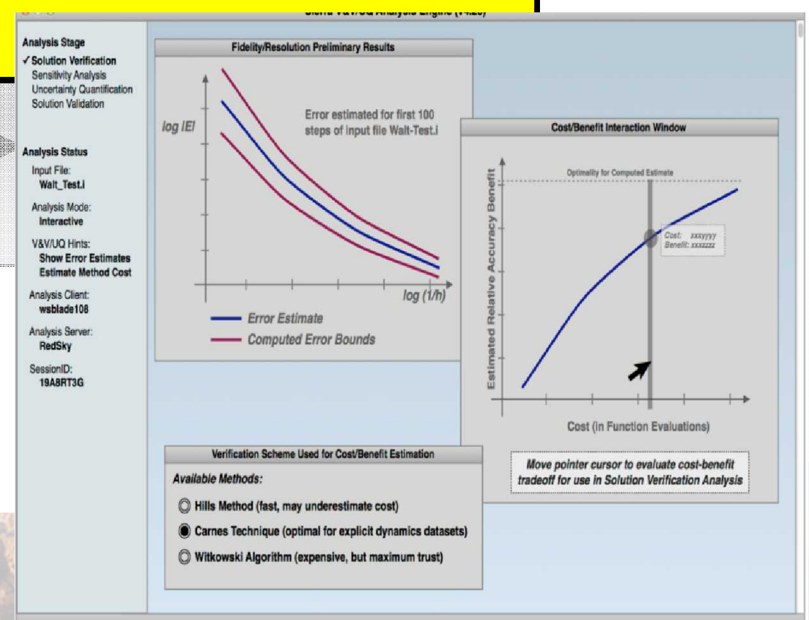
Intrinsic V&V Process: Solution Verification

- Tools to enable analysts to better and more easily perform spatial and temporal discretization studies
- Tools to enable analysts to better and more easily perform solution error and uncertainty estimates

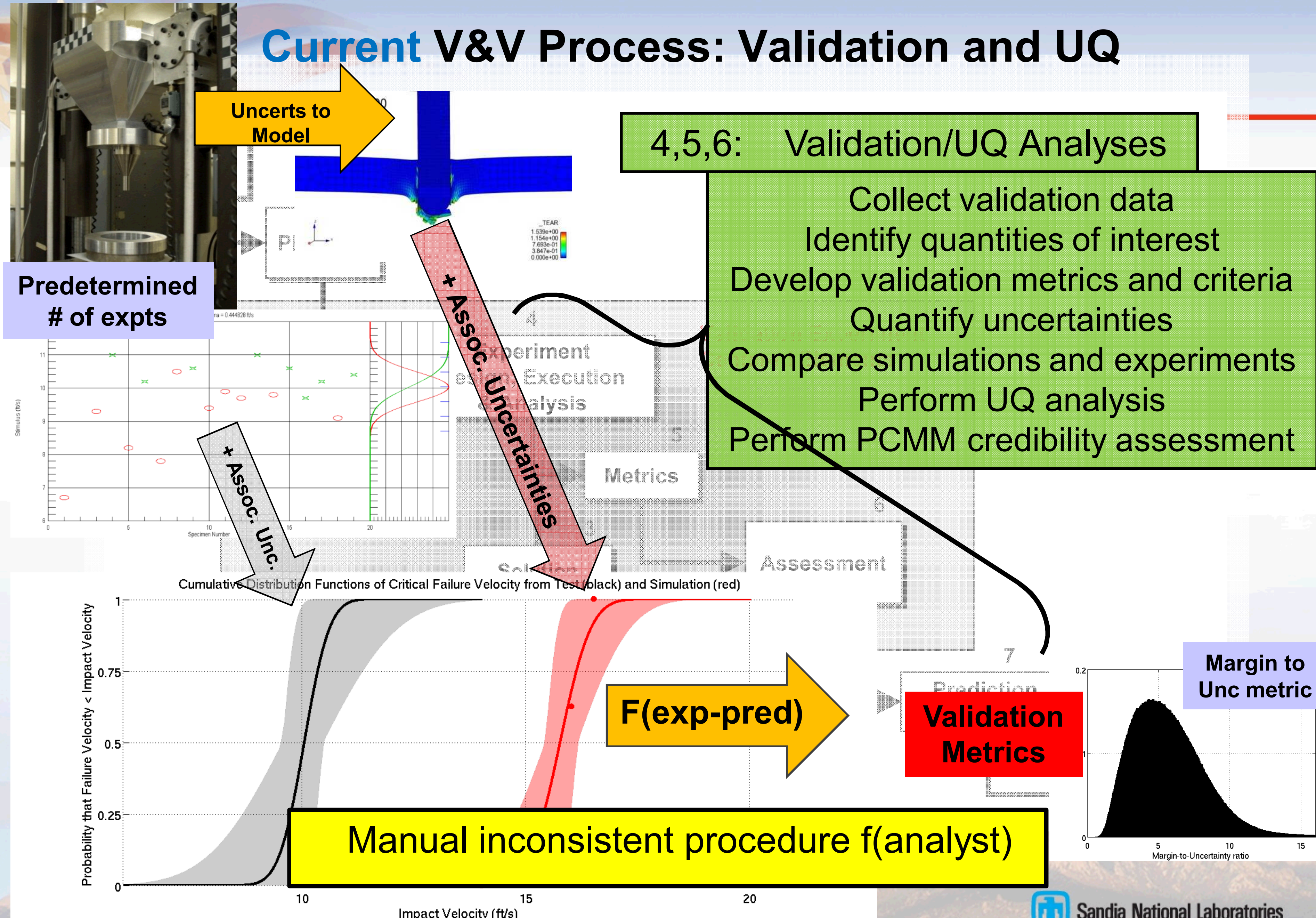
→Develop cost benefit resource estimation tool → minimize errors for given resource budget

→Recommended analysis parameters extracted from past application database → adaptive mesh refinement based on error estimate and estimated rate of convergence





Current V&V Process: Validation and UQ



Sandia National Laboratories

V&V/UQ Process
Integrated, Iterative Assessments (SA, Val, UQ)

Current Process

- Manual scripting to couple Dakota and ASC codes *per platform*
- Analyst must create input and output filters between Dakota and the analysis code
- Analysts must set up Dakota input files and parameterize Sierra input file
- Analysts should verify algorithm parameter selections

Future Process

- GUI-driven iterative analysis loop development in workbench (no reqt to know input file syntaxes)
- Automated code integration *per platform* (no human-in-the-loop errors; less startup time)
- Selectable input parameters, responses and output filters (no need to parameterized input files)

→ V&V/UQ capabilities embedded within the analysis codes and used internally

→ VALTS for validation verification (incl exptl data)

→ Links to past application for recommended V&V/UQ practices

→ Links to material models and characterization library with uncertainties

→ Based on reqts and past unc recommend exptl design

→ Optimal design based on associated unc and design reqts

Summary of future V&V/UQ Analysis → Intrinsic V&V

Short term:

- Must continue to change the philosophical usage of V&V/UQ from a mandated, after-thought, box-checking exercise to a necessary and valuable inseparable aspect of CompSim during planning and for credibility assessment evidence support.
- Enable the analyst with tools, consistent methodologies and procedures

  →consistent credibility and supporting evidence

Long term:

- Fully integrated, inseparable CompSim and V&V/UQ with link to archival database

→ Robust to analyst opinion (novice or seasoned)

→ Immune to changing demands on CompSim

→ Self-generating Validation and Qualification Plans *reverse engineered* based on project requirements, constraints and associated uncertainties

Sandia National Laboratories

Cost-Benefit for all IVV Tasks
(3 year time horizon)

IV&V Tasks:
Gen
CVER
SVER
VAL
SA/UQ

LT (Cost)	GEN – 3C CVER – 3B SVER - 1C VAL – 3B SA/UQ – 3B	GEN – 3D GEN – 3F SVER – 2C CVER – 3C CVER – 3D VAL – 2B VAL – 2D	VAL – 1B VAL – 1C SA/UQ – 3A GEN – 2C GEN – 3B
1.0	GEN – 3A GEN – 3E SVER – 2D SVER – 3B SVER – 3C SA/UQ – 1B SA/UQ – 2B SA/UQ – 2C	VAL – 2C SA/UQ – 1A SA/UQ – 1D SA/UQ – 2D SA/UQ – 2F	SA/UQ – 3C
0.5	GEN – 2D CVER – 1A CVER – 2C CVER – 4A CVER – 4D SVER – 2B SVER – 4A SVER – 4B VAL – 2E VAL – 3C VAL – 3D SA/UQ – 1C SA/UQ – 2A	CVER – 2F SVER – 1D SA/UQ – 2D SA/UQ – 2E	
0.25	GEN – 1A GEN – 1B GEN – 1C GEN – 1D GEN – 1E GEN – 1F GEN – 2A GEN – 2B CVER – 1C CVER – 2A CVER – 2B CVER – 3A SVER – 1A SVER – 2A SVER – 3A VAL – 1A VAL – 2A VAL – 3A	SVER – 1B CVER – 1B CVER – 2E CVER – 4C	CVER – 2D CVER – 2G CVER – 4B VAL – 1D
	1 (Highest Benefit)	2	3

IV&V Team:
V&V experts
Sierra POs
Dakota Rep
+ SNL/CA

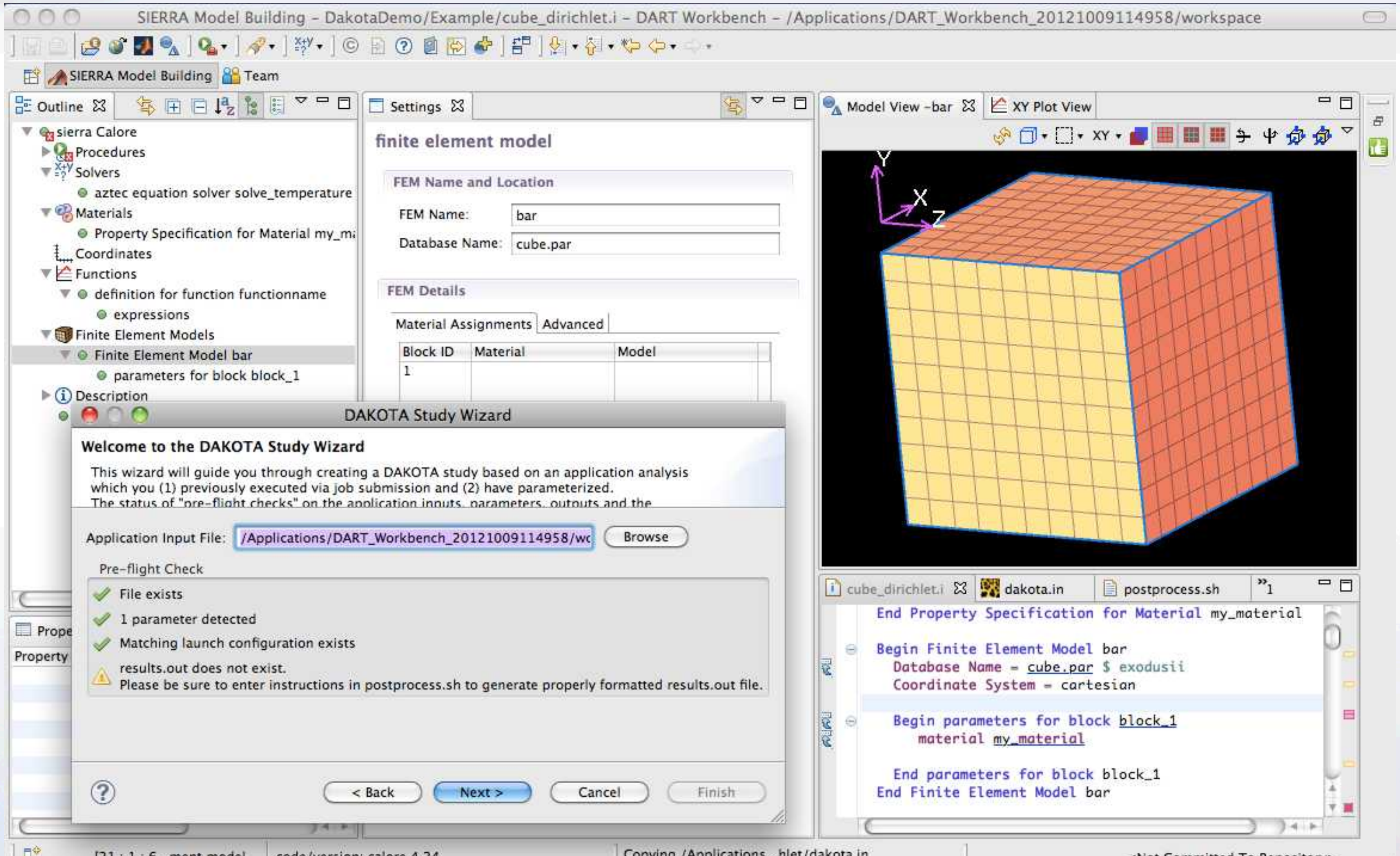
Sandia National Laboratories

Cost – Prioritization for Benefit=1 Tasks

High Cost			GEN – 3C VAL – 3B	SVER - 1C	SVER – 2D SVER – 3C	SVER – 3B		SA/UQ – 3B	GEN – 2D
		SA/UQ – 1B CVER – 2C CVER – 4D	CVER – 3B CVER – 1A GEN – 1A	SA/UQ – 2B SA/UQ – 2C VAL – 2E	GEN – 3E SA/UQ – 1C VAL – 3C	SA/UQ – 2A	GEN – 3A SVER – 4A		
Low Cost	CVER – 3A	GEN – 2A GEN – 2B CVER – 2B	GEN – 1B GEN – 1C GEN – 1D GEN – 1E GEN – 1F SVER – 2A	SVER – 1A CVER – 1C	CVER – 4A VAL – 3A CVER – 2A SVER – 3A	VAL – 3D VAL – 1A VAL – 2A	SVER – 2B	SVER – 4B	
	High Priority								Low Priority

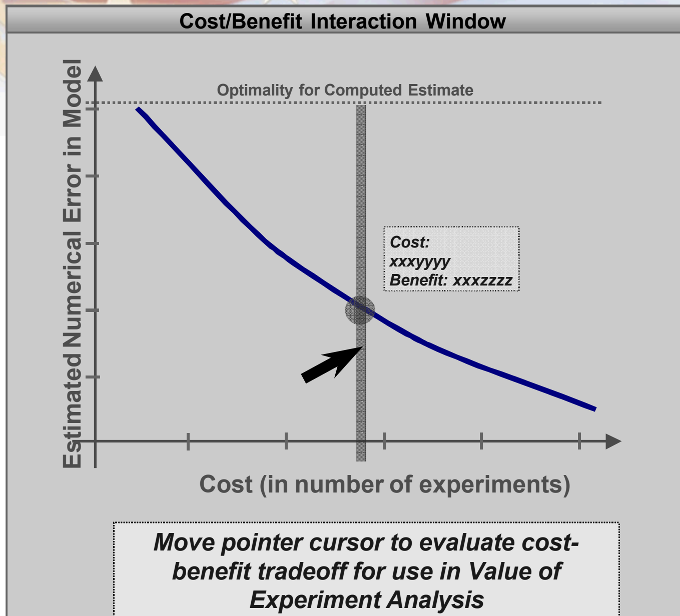
Extra Slides

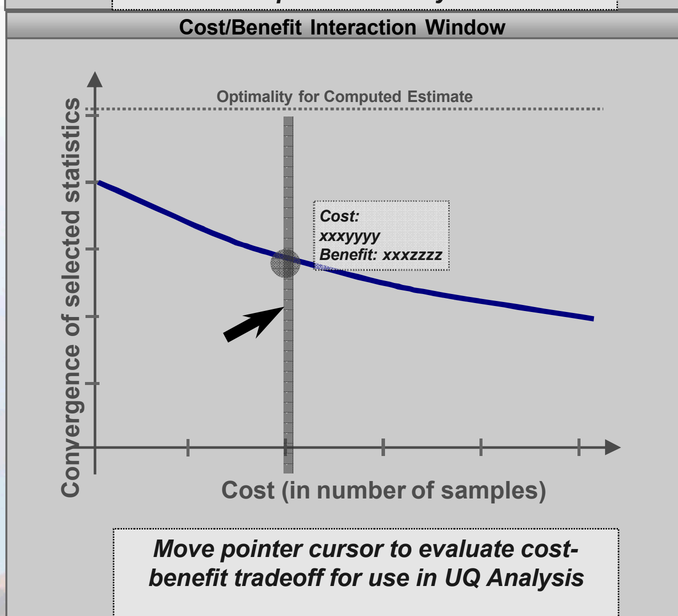
Tools Integration into Workbench



Integrate tools developed by the DART Workbench team and JAGUAR/DAKOTA into the CompSimUI that enable users to build and execute V&V/UQ analyses.

Intrinsic V&V Process: Validation/UQ





4,5,6,7: Validation/UQ Process

Collect validation data
Identify quantities of interest
Develop validation metrics and criteria
Quantify uncertainties
Compare simulations and experiments
Perform UQ analysis
Perform PCMM credibility assessment

- Develop cost benefit resource estimation tool → min errors for given resource budget
- More standard procedures/methodologies
- Develop VALTS FCT
- Create/Document VALTS with integrated exptl data
- Provide tools for validation and UQ workflow
- Library of validation metrics and other post-processing
- Std roll-up methodology (aggregate and propagate unc)
- Prescribed methods to handle separated uncertainties (epistemic, aleatory, numerical, parameter,…)