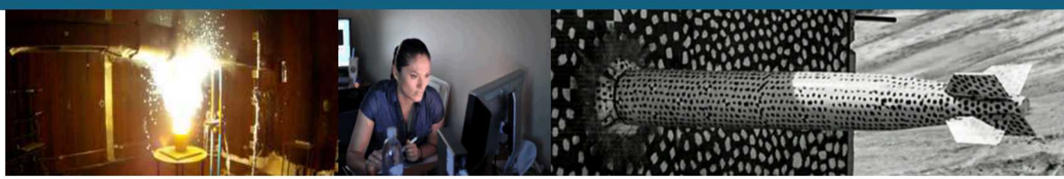
SAND2020-8968PE

Introduction to Verification, Validation (V&V), Uncertainty Quantification (UQ), and Credibility Processes



PRESENTED BY

Aubrey Eckert (1544)

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Credibility Evidence for Computational Simulation Predictions

How do we demonstrate that **predictions** derived from computational simulations are **credible**?

Expert judgement, I have been doing this for 50 years!

Credibility Evidence for Computational Simulation Predictions

How do we demonstrate that **predictions** derived from computational simulations are **credible**?

Expert judgement, I have been doing this for 50 years!

I ran the highest fidelity simulation on the best and biggest computer out there!

Credibility Evidence for Computational Simulation Predictions

How do we demonstrate that **predictions** derived from computational simulations are **credible**?

Expert judgement, I have been doing this for 50 years!

I ran the highest fidelity simulation on the best and biggest computer out there!

The deliverable is due today, so it better be credible!

Credibility Evidence for Computational Simulation Predictions

How do we demonstrate that **predictions** derived from computational simulations are **credible**?

Expert judgement, I have been doing this for 50 years!

I ran the highest fidelity simulation on the best and biggest computer out there!

The deliverable is due today, so it better be credible!

We used the same process we have always used, we have never been wrong before!

Credibility Evidence for Computational Simulation Predictions

How do we demonstrate that **predictions** derived from computational simulations are **credible**?

Expert judgement, I have been doing this for 50 years!

I ran the highest fidelity simulation on the best and biggest computer out there!

The deliverable is due today, so it better be credible!

We used the same process we have always used, we have never been wrong before!

We built conservatism and plenty of margin into all of our calculations!

Credibility Evidence for Computational Simulation Predictions

How do we demonstrate that **predictions** derived from computational simulations are **credible**?

Expert judgement, I have been doing this for 50 years!

I ran the highest fidelity simulation on the best and biggest computer out there!

The deliverable is due today, so it better be credible!

We used the same process we have always used, we have never been wrong before!

We built conservatism and plenty of margin into all of our calculations!

Look, my presentation has a cool video!

Credibility Evidence for Computational Simulation Predictions

How do we demonstrate that **predictions** derived from computational simulations are **credible**?

Expert judgement, I have
been doing this for 50

I ran the highest fidelity

Although aspects of these assertions may lend a certain level of credibility to analyses, these assertions **cannot stand alone** as the only **credibility evidence** to support a computational simulation prediction, particularly in a **high consequence** environment

today, so it better be

We used the same

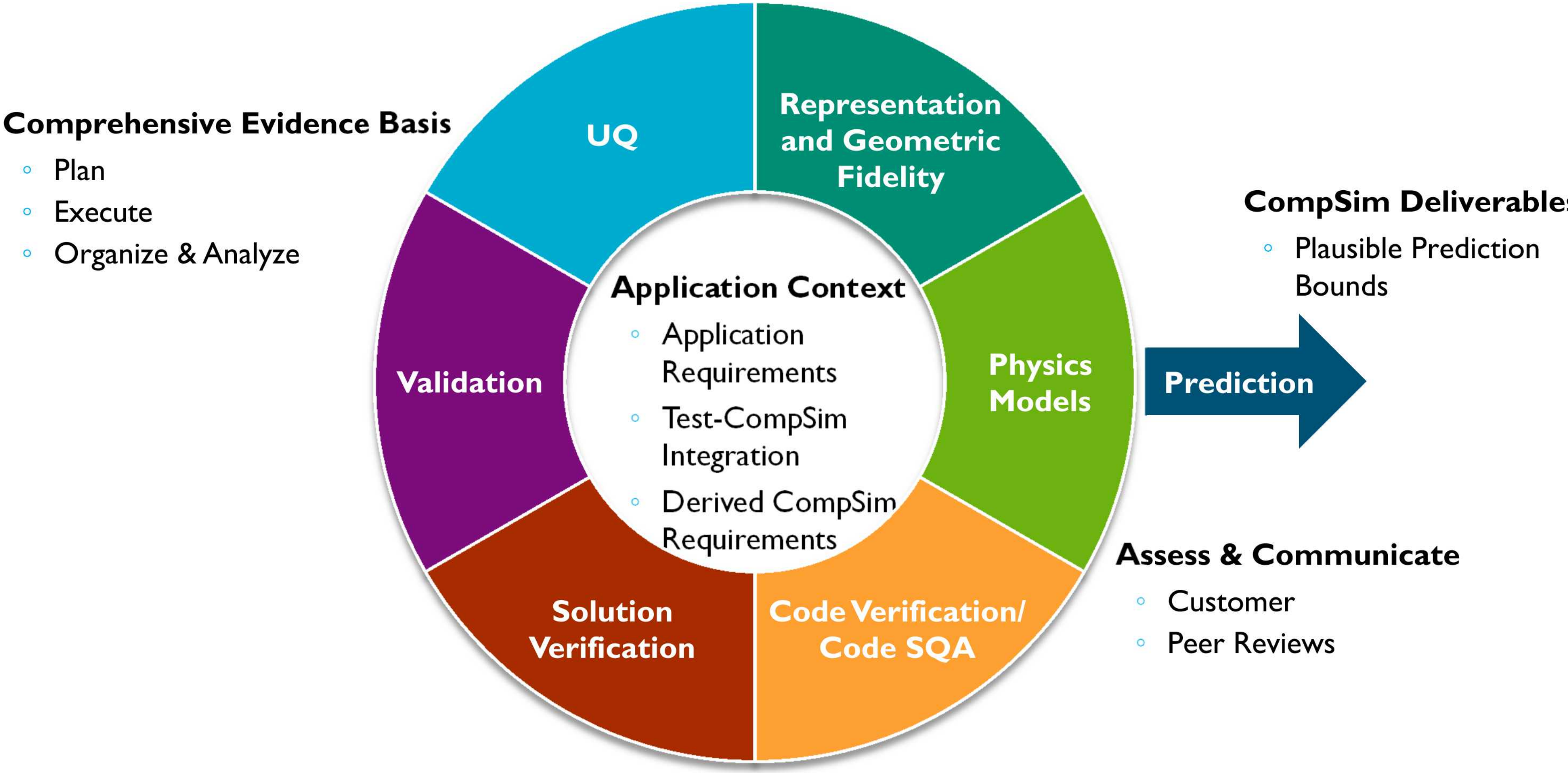
The **computational simulation credibility process** seeks to provide a documented, consistent, and repeatable process for assembling a **comprehensive credibility evidence package** to support computational simulation predictions

all of our calculations!

Look, my
presentation has
a cool video!

Credibility Evidence for Computational Simulation Predictions

- The computational simulation (CompSim) **credibility process** assembles and documents **evidence** to ascertain and communicate the **believability** of **predictions** that are produced from computational simulations.

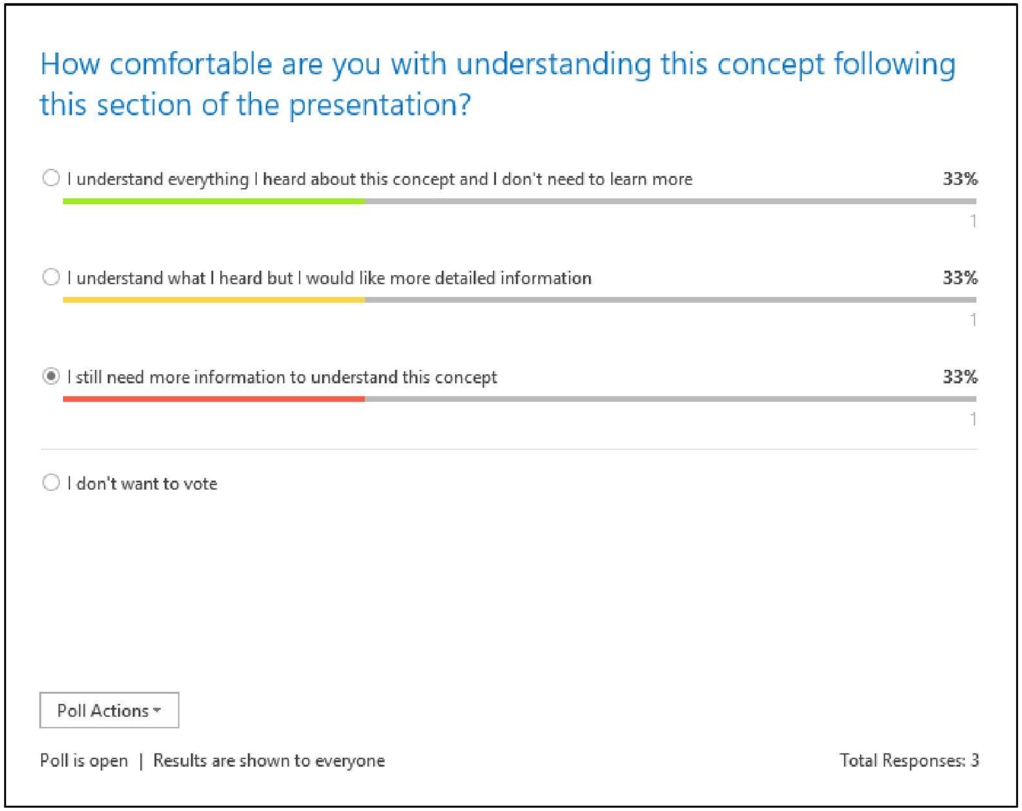


Goal & Structure of this Training

- Goal
  - The purpose of this short course is to **introduce** V&V/UQ/Credibility process concepts, methods, and tools that have been developed for CompSim at Sandia.
- Structure
  - This course will be a **2-hour overview** of V&V/UQ/Credibility processes for CompSim at Sandia. Additional information, including links the full ESP700 class, will be provided if more detail is needed.
- Feel free to interrupt, ask questions, and let me know where more information is needed – this will be more productive if it is a **dialogue** rather than a 1-sided presentation

Real-time Feedback

- We will be gathering **feedback** during the presentation to help us identify areas where additional clarity is needed.
  - After concepts or methods are presented, we will use the scale below to take a poll with entries in the skype chat window.
  - This is not intended to give you a grade – it will help us improve our delivery and materials and will inform future deep-dives.



Introduction to CompSim Credibility at Sandia

Motivation and Historical Perspective

- Important Motivations
  - The nation is making **million/billion dollar decisions** that are strongly influenced by CompSim
    - Weapon life extensions
    - Facility/infrastructure protection upgrades
    - Spacecraft launches
    - More
  - How do we build/demonstrate **confidence** in our CompSim results?
- V&V/UQ/Credibility is expected, but is not always well understood, by decision makers.
  - V&V is, in a nutshell, all about putting "correct" math methods and physics models into our codes.
  - We are expected to produce "correct" codes and models which leads to "correct" results.
- In the past V&V/UQ/Credibility was an **afterthought** if thought of it at all and was sometimes considered to be a **nuisance**.

Motivation and Historical Perspective

- What is different now?
  - CompSim is different now than 10-20-30 years ago (e.g., auto industry, aircraft industry, nuclear weapons (NW) industry)
    - We are already making million/billion dollar decisions that are heavily influenced by CompSim.
  - "Before I spend \$M/\$B on a decision, I want **evidence** of the correctness of your CompSim results."

- Issues:
  - Correctness is expected or implied, but is **not innate** and **requires extra effort** to provide quantitative evidence (via V&V/UQ/Credibility processes)
  - Due to **resource constraints**, you can't V&V every aspect of a code/model/project, run the perfect UQ study for every analysis, or provide a comprehensive collection of credibility evidence for every calculation
  - It's hard to **retrofit** V&V/UQ/Credibility activities into a study that is already completed.

Motivation and Historical Perspective

- V&V is **not palatable for its own sake**.
  - Decision makers don't care about the rate of convergence of an iterative mathematical method, or percent line coverage of tests.

- For million/billion dollar issues, decision makers do care that you got the right answer and they **expect a technical pedigree** (provenance) for your work.

- V&V is palatable when it is included as an **aid to decision making**.
  - V&V provides supporting evidence (provenance) to sensitivity analysis and UQ results on relevant technical/financial issues.
  - V&V/UQ/Credibility activities provide evidence that help to **buy down the risk** incurred by basing decisions on CompSim predictions
    - Utilizing testing for these decisions is not without similar risks. Experimental credibility evidence and activities are also important.

Where is SNL Now W.R.T. CompSim & V&V/UQ/Credibility?

- SNL NW **mission drivers**:
  - Annual assessment & certification that all weapon types are safe, secure & reliable
- Few/no tests at the full system level; few/some/no tests at subsystem/component level:
  - Not allowed (radiation effects tests)
  - Too expensive (crash impact tests)
  - Too environmentally unfriendly (fuel/propellant fire tests)
  - Too few units available (annual surveillance)
- In ~1996, Sandia entered the Stockpile Stewardship Program to develop CompSim tools to:
  - Aid in decision making in the absence/reduction of test data, and
  - Improve the technical basis (i.e., understanding) of the basic physical processes that dictate weapon performance in all environments.
- In ~2007, Sandia NW Engineering community embraced CompSim (particularly high-fidelity CompSim) as an integral part of the NW design/analysis/qualification process.
  - Sandia NW Engineering is putting in place the policies, procedures, and peer reviews that essentially **mandate** V&V on all significant CompSim studies.

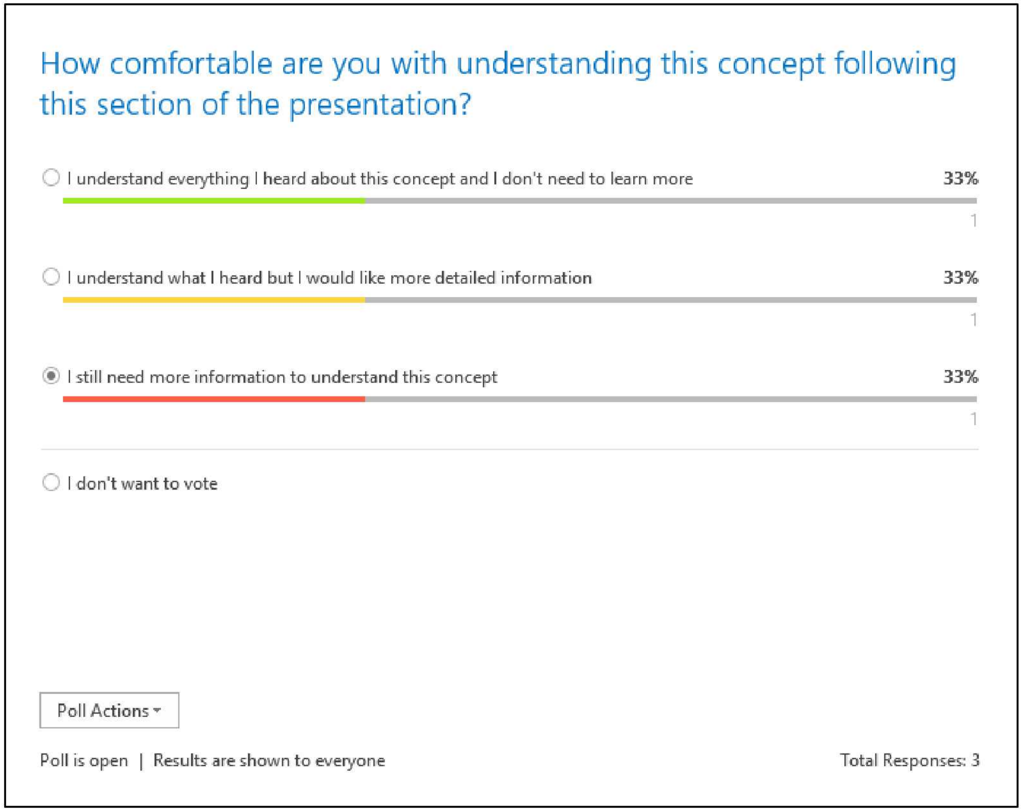
Sandia's V&V/UQ/Credibility Community

- Sandia's V&V/UQ/Credibility process experts **develop methods and tools** and **deploy** these tools to support the collection of credibility evidence for CompSim predictions.
  - 1544 deploys staff members as **embedded V&V partners** who develop **strategies** for collecting credibility evidence and **execute** credibility activities.
  - Other partners and departments also support credibility activities:
    - California V&V/UQ departments and partners (8750)
    - Statistics department team members (6673)
    - Dakota team (1463)
  - Education materials and capabilities are made available so that staff outside of these departments can learn about and utilize V&V/UQ/Credibility tools themselves

- The most **effective** teams supporting credibility activities include the subject matter experts, code developers, analysts, experimentalists, and leaders who are responsible for the analyses and predictions of interest.
  - V&V partners cannot perform credibility activities alone.
  - Level of hands-on involvement is based on many factors.

Real-time Feedback

- Concepts:
  - Motivation, history, and current status of CompSim credibility at Sandia



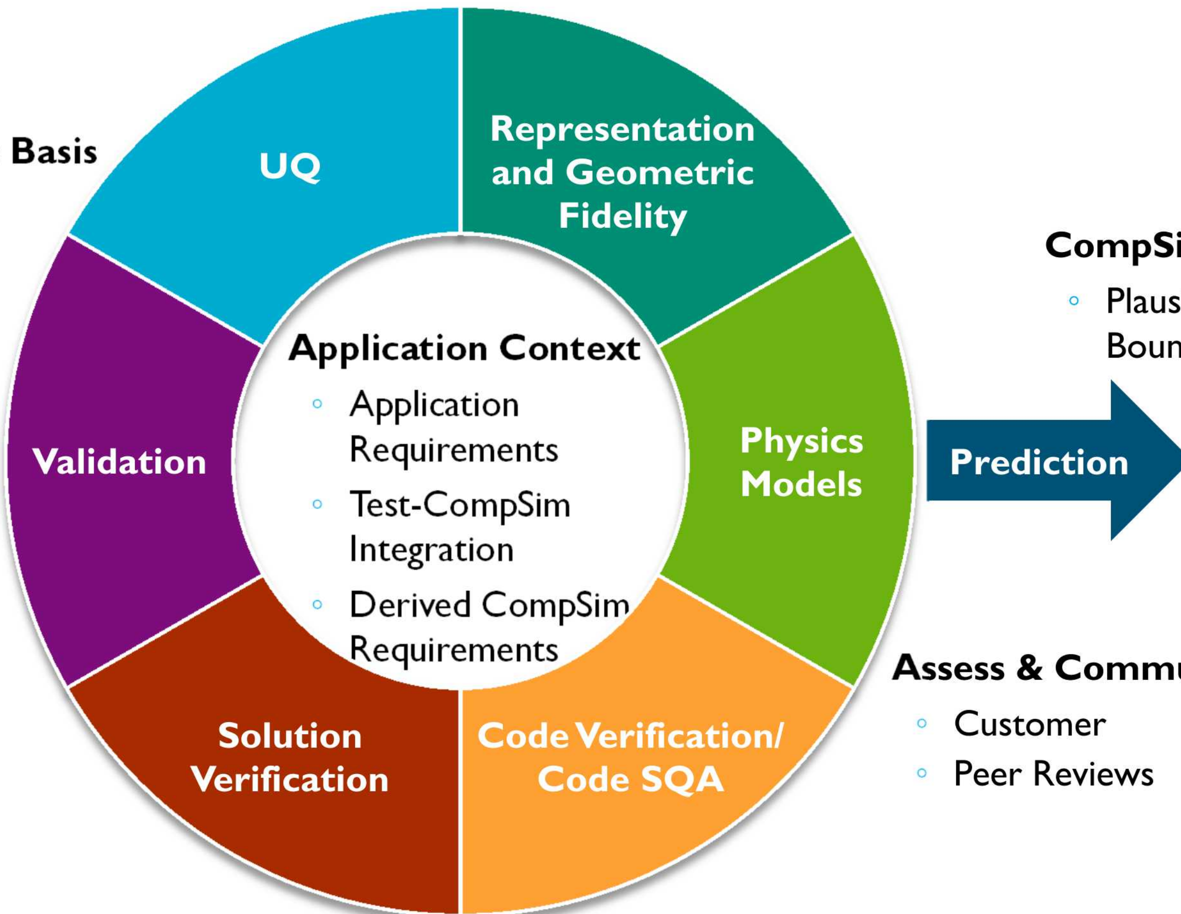
Integrating V&V/UQ/Credibility into CompSim

Credibility Evidence for Computational Simulation Predictions

- The computational simulation (CompSim) **credibility process** assembles and documents **evidence** to ascertain and communicate the **believability** of **predictions** that are produced from computational simulations.

Comprehensive Evidence Basis

- Plan
- Execute
- Organize & Analyze



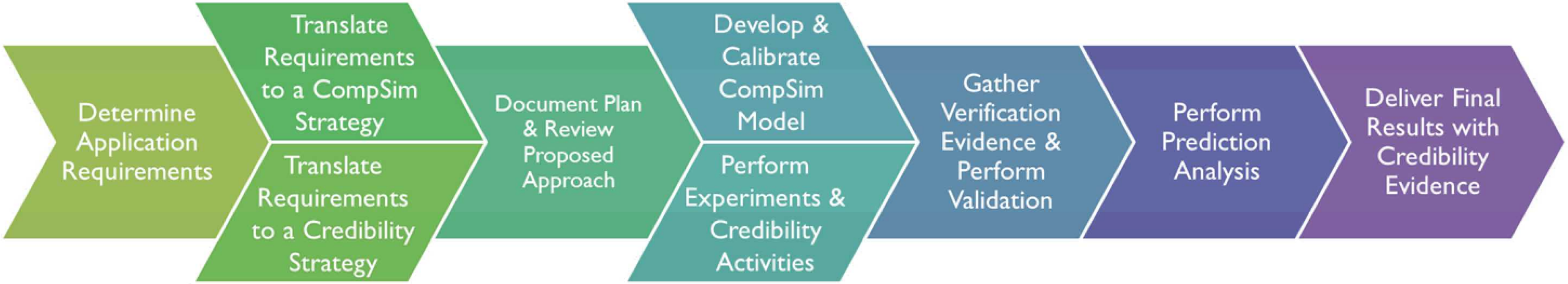
CompSim Deliverables

- Plausible Prediction Bounds

Assess & Communicate

- Customer
- Peer Reviews

Integrating V&V/UQ/Credibility into the CompSim Workflow



- The V&V/UQ CompSim workflow seeks to present the **end-to-end process** for planning for, developing, and presenting **credibility evidence** in support of modeling and experimental efforts that are used to develop CompSim predictions.

- In current practice, these practices can sometimes be **separated** from the general planning and activities supporting the development of a CompSim prediction.

- Methods and tools for more closely **integrating** V&V/UQ/Credibility processes with CompSim analyses continue to be developed.

- **Agile Credibility** seeks to provide V&V/UQ/Credibility planning and support in an efficient manner depending on the type of analysis and resource constraints.
  - This is an important goal for future V&V/UQ/Credibility development.

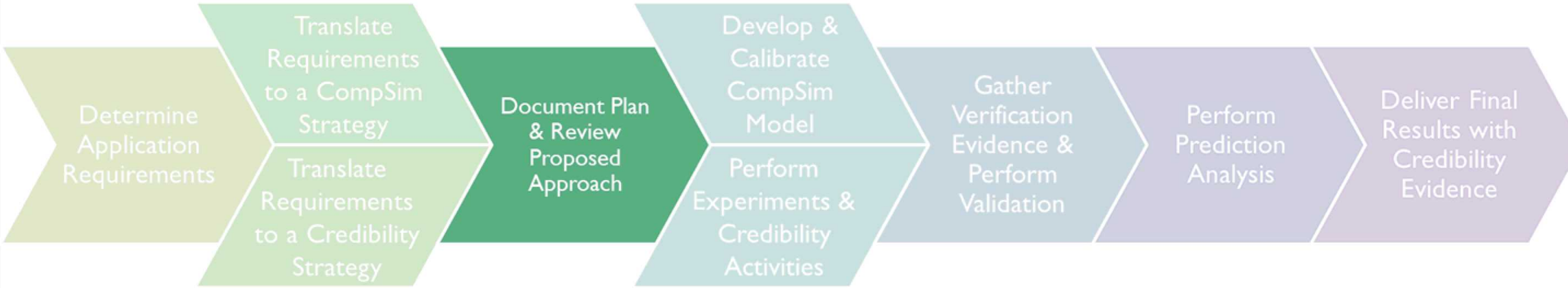
Integrating V&V/UQ/Credibility into the CompSim Workflow

- Defining and understanding **requirements** is a critical component of planning both analysis and V&V/UQ/Credibility activities
- Defining requirements includes:
  - Gathering requirements for the analysis
  - Defining the response, environment, and **quantities of interest (QoIs)**
  - Understanding the project constraints in terms of time, budget, computational capabilities
- Examples:
  - X quantity of interest cannot exceed a value of Y
  - The 99th percentile of X quantity of interest should fall below a value of Y with 95% confidence

Integrating V&V/UQ/Credibility into the CompSim Workflow

- Translating requirements into an application **strategy** sets the stage for executing V&V/UQ/Credibility activities as the analysis progresses
- This includes:
  - Identifying physics, math model, code capability, computational, and experimental requirements
    - The **Phenomena Identification and Ranking Table (PIRT)** tool supports this activity
  - Mapping these requirements into a plan for developing credibility evidence
    - The **Predictive Capability Maturity Model (PCMM)** is a useful tool that supports this translation
- Examples:
  - Completion of a design analysis may require model development and deterministic analyses that will aid in a design selection – design of experiments, extensive validation activities, and detailed uncertainty quantification may not be needed
  - Developing predictions in support of qualification activities requires extensive V&V/UQ/Credibility activities including verification, validation, and uncertainty quantification with close teaming amongst analysts, experimentalists, and V&V/UQ partners

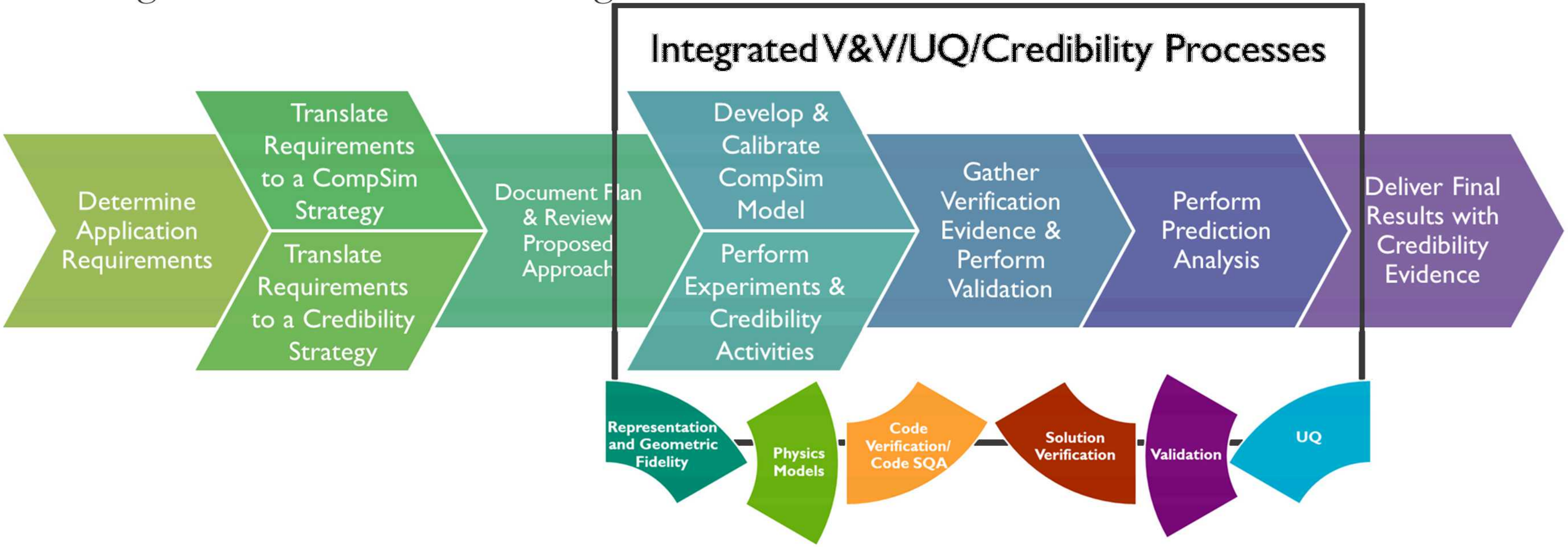
Integrating V&V/UQ/Credibility into the CompSim Workflow

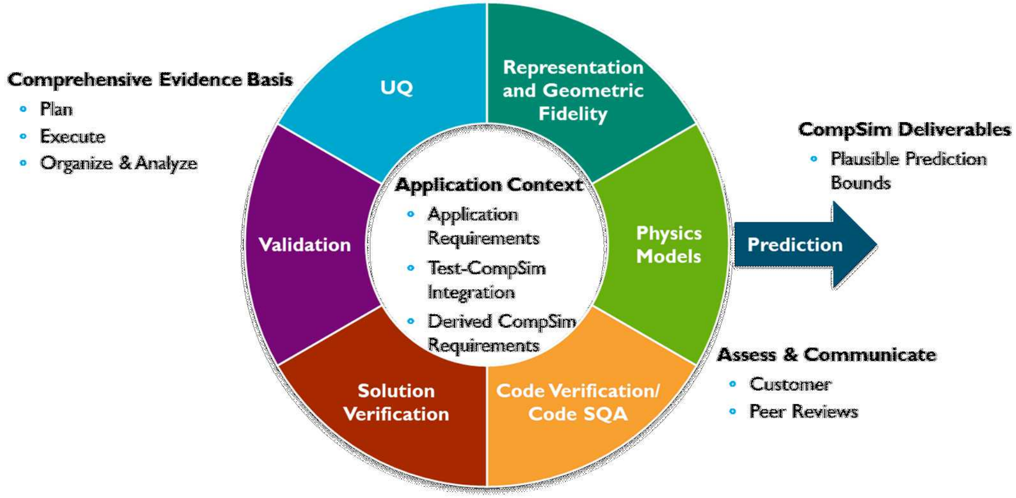


- **Documenting** the plan and reviewing it ensures that everyone is on the same page before work is executed
- Work should be reviewed to ensure that plans respect project constraints
- **Peer review** should take place at varying levels depending on application requirements
- Iteration on planning activities may be required to ensure that strategy aligns with requirements and that gaps are mitigated

Integrating V&V/UQ/Credibility into the CompSim Workflow

- We will go through the detailed V&V/UQ/Credibility components of the workflow using the PCMM elements as a guide

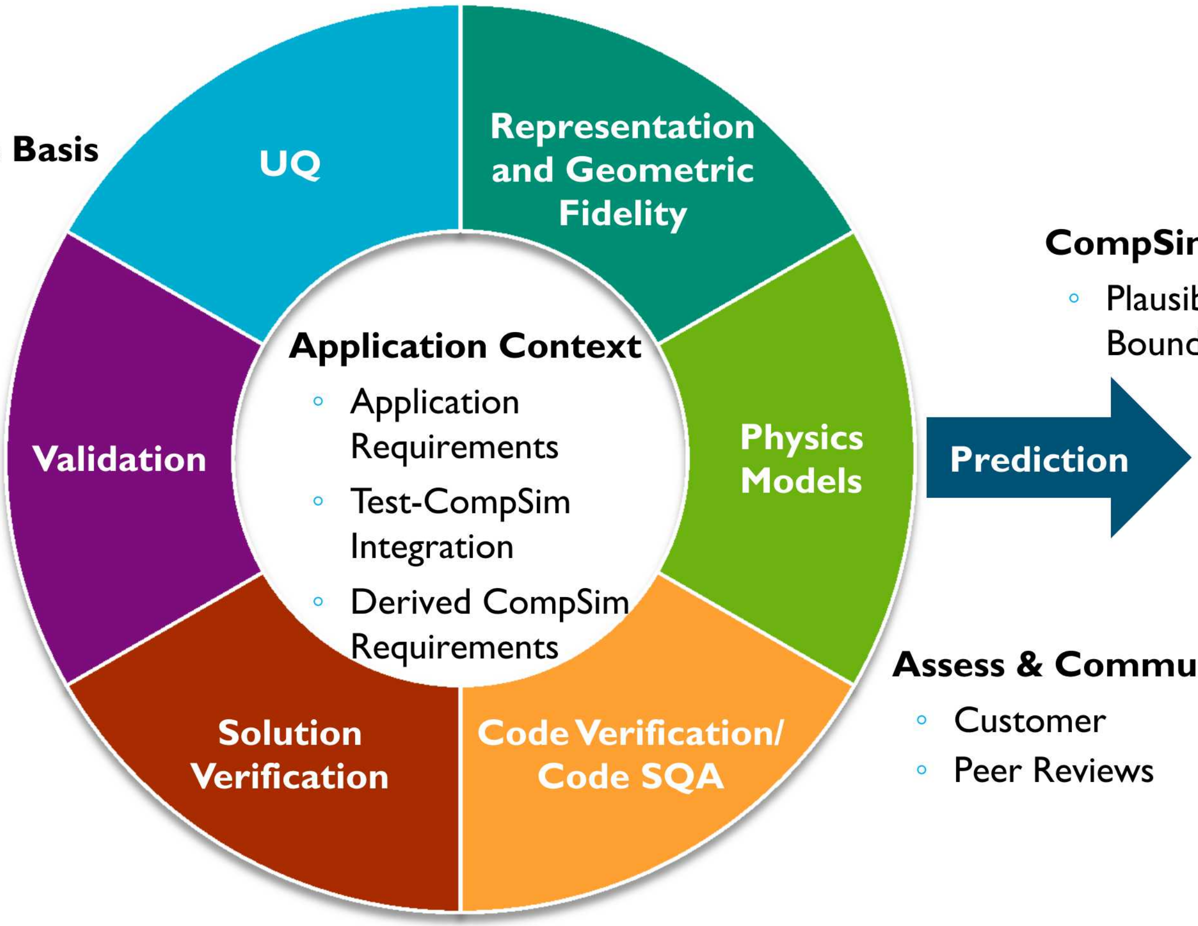




CompSim Credibility Process

- The process of assembling and documenting **evidence** to ascertain and communicate the **believability** of **predictions** that are produced from computational simulations.
  - **Application Context**
  - **Evidence-Related Elements**
  - **Prediction Issues**
  - **Key Gaps and Potential Paths Forward**

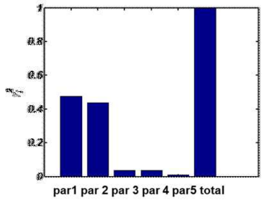
Comprehensive Evidence Basis
- Plan
- Execute
- Organize & Analyze

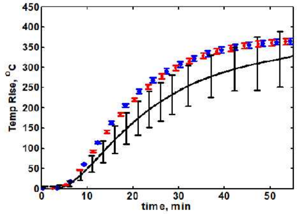


CompSim Deliverables
- Plausible Prediction Bounds

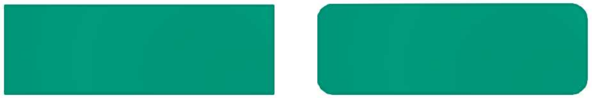
Assess & Communicate
- Customer
- Peer Reviews

CompSim Credibility Elements

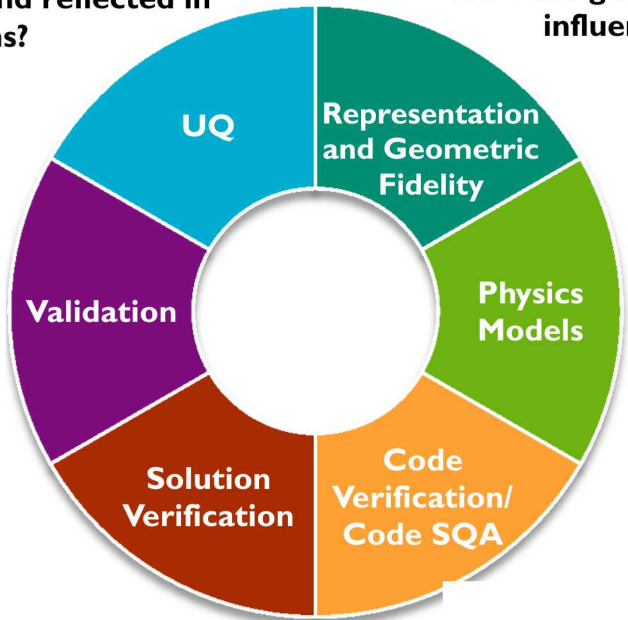


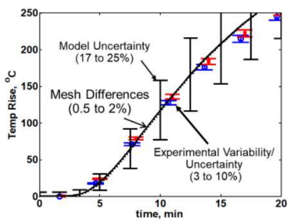


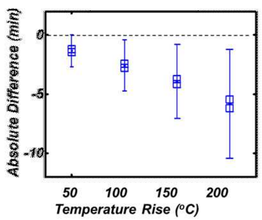
How are uncertainties assessed and reflected in simulation predictions?



How are geometric feature simplifications influencing simulation results?





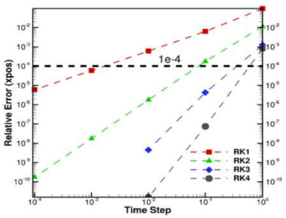


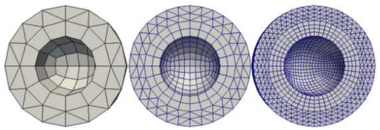
What is the discrepancy between simulation and experiments?

PIRT

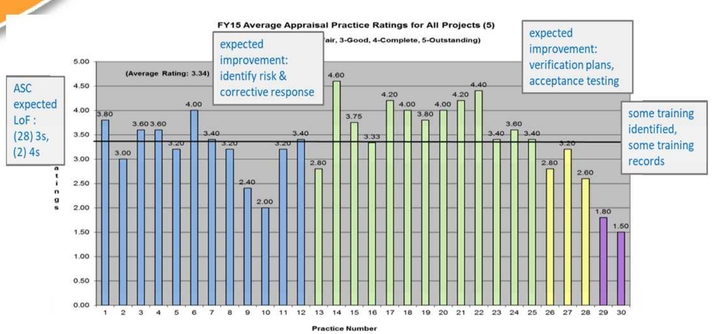
Phenomena	Importance	Adequacy for Intended Use: Math Model	Code	Validation	Model Parameter
Phenomena 1	H	H	M	L	L
Phenomena 2	M	H	M	L	L
Phenomena 3	L	H	M	L	L

Are important physics models adequate? Key gaps mitigated?





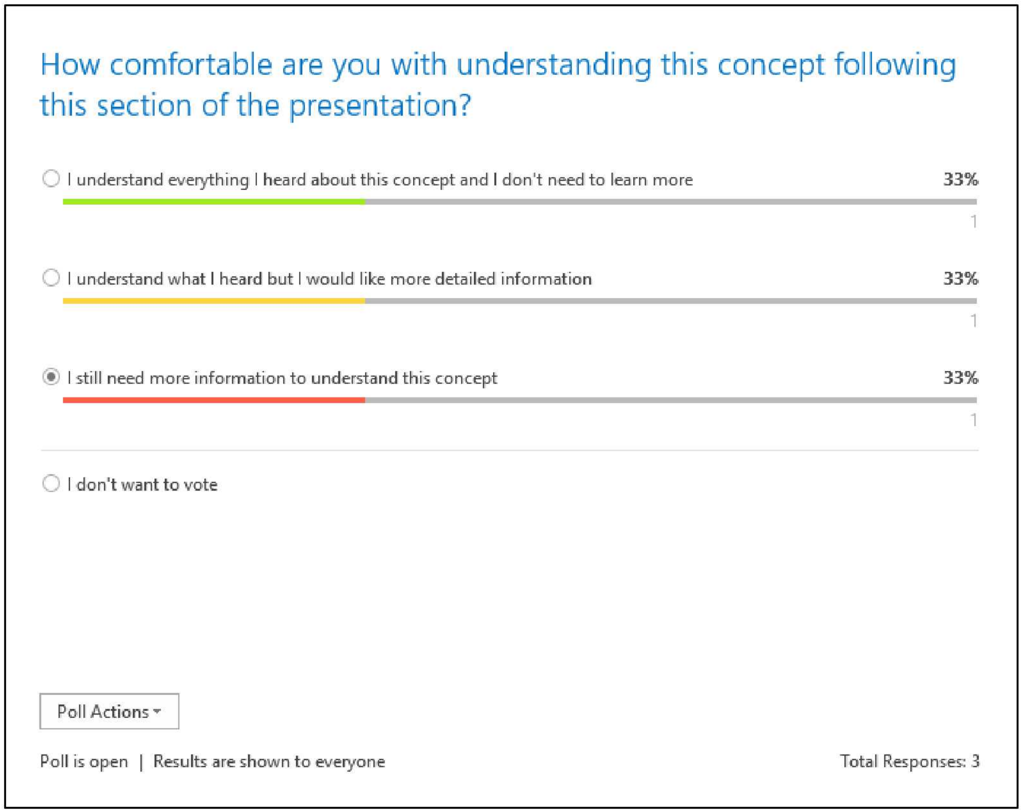
How do numerical solution or human errors affect simulation results?



What is the evidence for code credibility?

Real-time Feedback

- Concepts:
  - Mapping Requirements to a V&V/UQ/Credibility strategy
  - Understanding how credibility evidence supports simulation predictions at a high level



Credibility Elements Overview

What is the PCMM?

The **Predictive Capability Maturity Model (PCMM)** is a multi-dimensional qualitative metric to facilitate discussion and communication of **credibility evidence**.

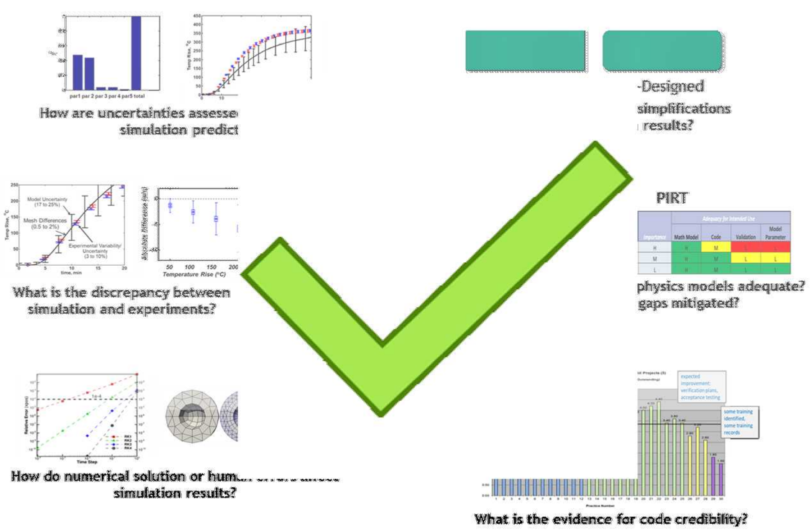
- Primary purposes:
  - Provide evidence to help determine **readiness** of modeling capabilities and simulation products for use in various applications and decisions
  - **Identify gaps** in the current credibility evidence for an application and **prioritize** additional activities
  - **Measure progress** of an integrated simulation effort **over the lifetime** of an analysis
- PCMM components:
  - **Elements** – the dimensions of the credibility evidence
  - **Maturity** – the state of the evidence and level of effort around each element
  - **Element criteria** – major features of the credibility evidence to consider for each element

Origins of the PCMM

- The PCMM was developed at Sandia National Laboratories
  - The need to develop a framework to **assess** CompSim analyses arose as **CompSim** became more **heavily relied upon** to design and assess the safety of engineered systems.
  - Sandia has deployed the PCMM across a wide variety of applications and physics disciplines
- The original PCMM has been **expanded** and **iterated upon** since its development
  - Iterations have increased the level of granularity for the PCMM elements
  - Method of deploying PCMM has changed through time and with lessons-learned

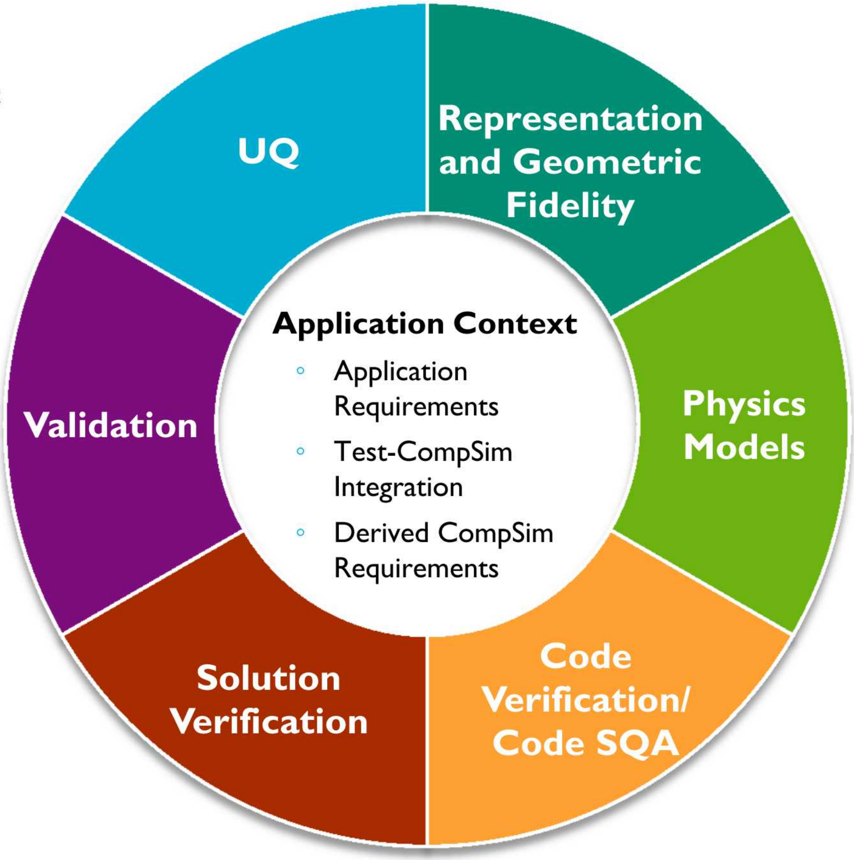
What the PCMM is and What the PCMM is Not

- The PCMM **IS**:
  - A planning tool to **highlight** and **prioritize** detailed V&V/UQ activities at an early stage of an analysis
  - A **communication tool** that *must* include a discussion of the **supporting evidence** to tell a credibility story
  - A tool for **informing risk** related to the use of modeling and simulation
- The PCMM is **NOT**:
  - An **absolute number** or a **score**
  - A mechanism for **criticizing** or **poking holes** in analysis credibility



What are the Outcomes of the PCMM?

- The PCMM is used to:
  1. Guide the **collection** of a **comprehensive** set of **credibility evidence**
  2. Organize the evidence to **communicate** the credibility story to decision makers
- The credibility evidence must exist before it can be evaluated
  - What evidence will be generated?
  - Will it tell a coherent story?
  - Will it be adequate?
  - If evidence does not exist, the PCMM will identify this as a gap
- The PCMM elements represent the **dimensions** of the evidence
  - Representation and Geometric Fidelity
  - Physics and Material Model Fidelity
  - Code Verification
  - Solution Verification
  - Validation
  - Uncertainty Quantification



Prerequisite Steps

- A subset of the team including the PCMM facilitator and team lead should meet to review **prerequisite** materials and questions.

- Prerequisite materials include:
  - Defining CompSim **objectives**
  - Determining **status** of modeling and V&V/UQ efforts
  - Completing a PIRT (Phenomena Identification and Ranking Table)

Objectives of the CompSim Activities

- Defining the **overall objectives** of the CompSim activities is important to the success of the PCMM.
- Understanding the **application requirements** that need to be met helps to determine the required level of credibility evidence that must be gathered.
- The PCMM begins with answering the following questions:
  - What is the **context** of the modeling activities?
  - Who are the **primary stakeholders** for this effort?
  - How will the **simulation outcomes** be used by **decision makers**?
  - What are the analysis **scenarios of interest**?
  - What are the **quantities of interest** (QoIs) and **prediction objectives**?
  - What are the **deliverables** and **timelines** for these activities?

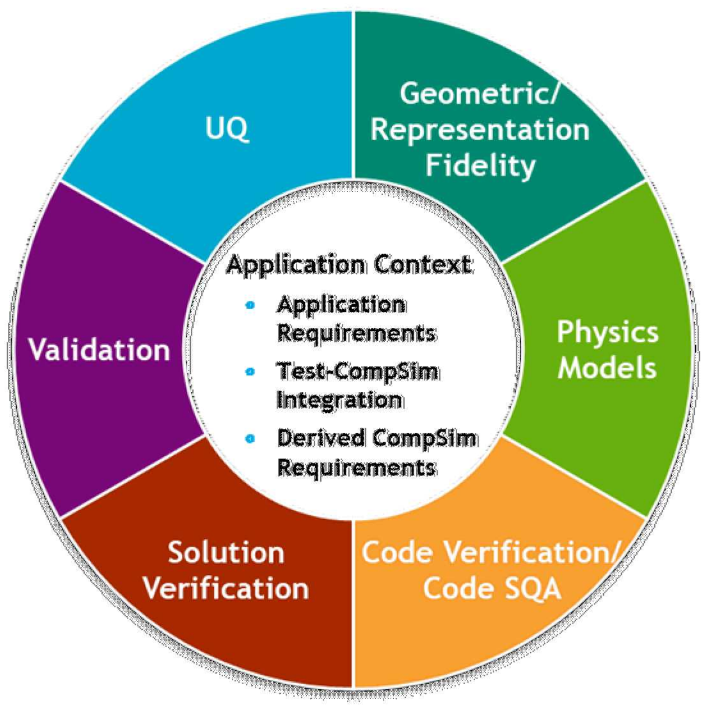
Status of Modeling and V&V/UQ Efforts

The following prerequisite steps and questions must be considered before the PCMM continues:

- Has a **PIRT** been conducted? If not, consider doing one first. If so, reference key high-level findings here.
- What is the **current stage** of the modeling effort for this application? (e.g., planning of activities, communication with stakeholders, etc.)
- What are the **goals** of this PCMM activity? (e.g., develop a V&V/UQ plan, develop a credibility story to communicate)

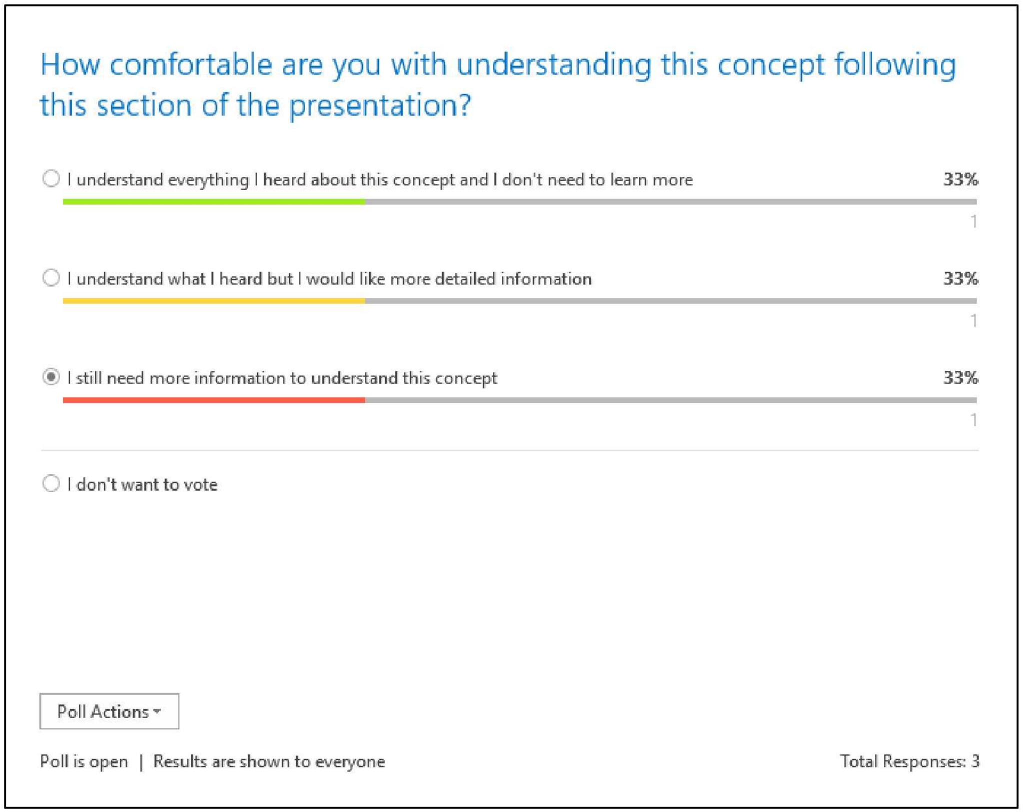
PCMM Elements Deep Dive

- Each PCMM element is divided into **sub-elements**
- PCMM **sub-elements** have been broken into a **series of questions** that provide detailed information related to the collection of credibility evidence.
- As the project team answers each question, **existing credibility evidence** and **gaps** in this credibility evidence will be identified.
- Discussions should include a relationship back to the **application context** and **re**



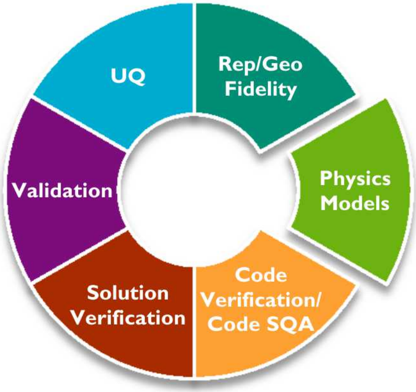
Real-time Feedback

- Concepts:
  - What PCMM is, why we use it, and what the intended outcomes are



Physics and Material Model Fidelity

Physics and Material Model Fidelity



Physics and Material Model Fidelity – "Are the important physics models adequate?

- The process of characterizing modeling completeness and adequacy for intended application.

➤Tool: PIRT

Phenomena	*Importance*	*Adequacy for Intended Use*			
		Math Model	**Code**	**Validation**	**Model Parameter**
Phenomena 1	H	H	M	L	L
Phenomena 2	M	H	M	L	L
Phenomena 3	L	H	M	L	L

Phenomena Identification and Ranking Table (PIRT) *often used at Sandia to support this element.*

PIRT: **P**henomena **I**dentification and **R**anking **T**able

- Define **key physical phenomena** and rank their importance for a particular quantity of interest
- Importance is relative to **quantity of interest** in the application scenario
- Assess **adequacy** and **gaps** in simulation capabilities and available data
- Adequacy of capabilities is relative to **intended use**
- Gaps are identified when adequacy scoring is below importance ranking

PIRT Steps

- Characterize completeness versus the PIRT
  - A PIRT should have already been completed for this analysis.
  - To what extent do the phenomena covered in the PIRT align with the major physics that are included in the application model, and are the same capabilities that were assessed in the PIRT applicable?

Step 1 | Step 2 | Step 3

Phenomena	Importance	Adequacy for Intended Use			
		Math Model	Code	Validation	Model Parameter
Phenomena 1	H	H	M	L	L
Phenomena 2	M	H	M	L	L
Phenomena 3	L	H	M	L	L

Step 4: Gap Assessment ← -2, -1, 0

Phenomena: Physical features or behaviors of an engineering analysis that are relevant to the intended application

Importance: Level of importance to quantity of interest in application scenario

Adequ… s,

If you were to model this problem from scratch, what physics would you need to include?

High, Medium, Low, Unknown

PIRT Steps

- Characterize completeness versus the PIRT
  - A PIRT should have already been completed for this analysis.
  - To what extent do the phenomena covered in the PIRT align with the major physics that are included in the application model, and are the same capabilities that were assessed in the PIRT applicable?

Step 1	*Step 2*	*Step 3*			
		Adequacy for Intended Use			
Phenomena	Importance	Math Model	Code	Validation	Model Parameter
Phenomena 1	H	H	M	L	L
Phenomena 2	M	H	M	L	L
Phenomena 3	L	H	M	L	L

Step 4: Gap Assessment ← -2, -1, 0

Phenomena: Physical features or behaviors of an engineering analysis that are relevant to the intended application

Importance: Level of importance to quantity of interest in application scenario

Adequacy: Ranking capabilities (mathematical models, material models, codebase, validation) for their intended use

High, Medium, Low, Unknown

PIRT Steps

- Characterize completeness versus the PIRT
  - A PIRT should have already been completed for this analysis.
  - To what extent do the phenomena covered in the PIRT align with the major physics that are included in the application model, and are the same capabilities that were assessed in the PIRT applicable?

Step 1	*Step 2*	*Step 3*			
		Adequacy for Intended Use			
Phenomena	*Importance*	Math Model	Code	Validation	Model Parameter
Phenomena 1	H	H	M	L	L
Phenomena 2	M	H	M	L	L
Phenomena 3	L	H	M	L	L

Step 4: Gap Assessment -2, -1, 0

Phenomena: Physical features or behaviors of an engineering analysis that are relevant to the intended application

Importance: Level of importance to quantity of interest in application scenario

Adequacy: Ranking capabilities (mathematical models, material models, codebase, validation) for their intended use

High, Medium, Low, Unknown

PIRT Steps

- Characterize completeness versus the PIRT
  - A PIRT should have already been completed for this analysis.
  - To what extent do the phenomena covered in the PIRT align with the major physics that are included in the application model, and are the same capabilities that were assessed in the PIRT applicable?

Step 1 ***Step 2*** ***Step 3***

Phenomena	Importance	Adequacy for Intended Use			
		Math Model	Code	Validation	Model Parameter
Phenomena 1	H	H	M	L	L
Phenomena 2	M	H	M	L	L
Phenomena 3	L	H	M	L	L

Step 4: Gap Assessment ← -2, -1, 0

Phenomena: Physical features or behaviors of an engineering analysis that are relevant to the intended application

Importance: Level of importance to quantity of interest in application scenario

Adequacy: Ranking capabilities (mathematical models, material models, codebase, validation) for their intended use

High, Medium, Low, Unknown

Physics and Material Model Fidelity Sub-elements

- Quantify model accuracy (i.e., separate effects model validation)
  - Which individual phenomena have specific validation comparisons?
  - How were the existing validation comparisons conducted (quantitative vs. qualitative), and how was experimental uncertainty/error in the test data incorporated?
  - Where does the validation data come from, and are the comparisons documented?

		Adequacy for Intended Use			
Phenomena	*Importance*	Math Model	Code	Validation	Model Parameter
Phenomena 1	H	H	M	L	L
Phenomena 2	M	H	M	L	L
Phenomena 3	L	H	M	L	L

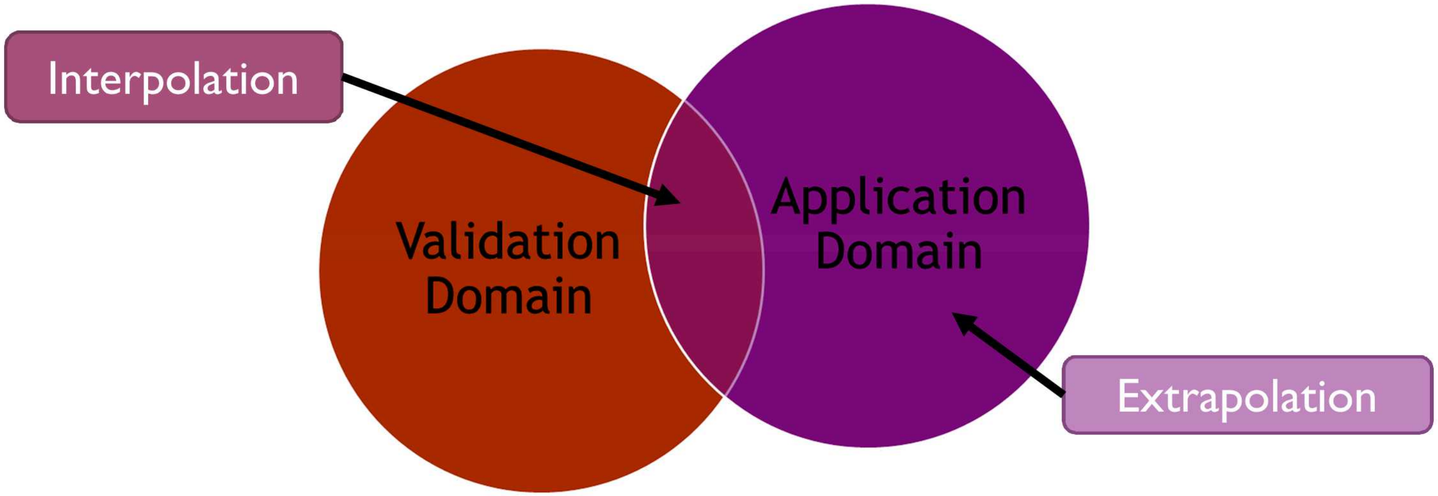
High: **Relevant test data** is available for the phenomenon, and **quantitative comparisons** have been made between the test data and the model outputs.

Medium: **Some relevant test data** is available for the phenomenon, but it has only been **qualitatively compared** with the model outputs or no comparison has been performed.

Low: **No relevant test data** is available for the phenomenon.

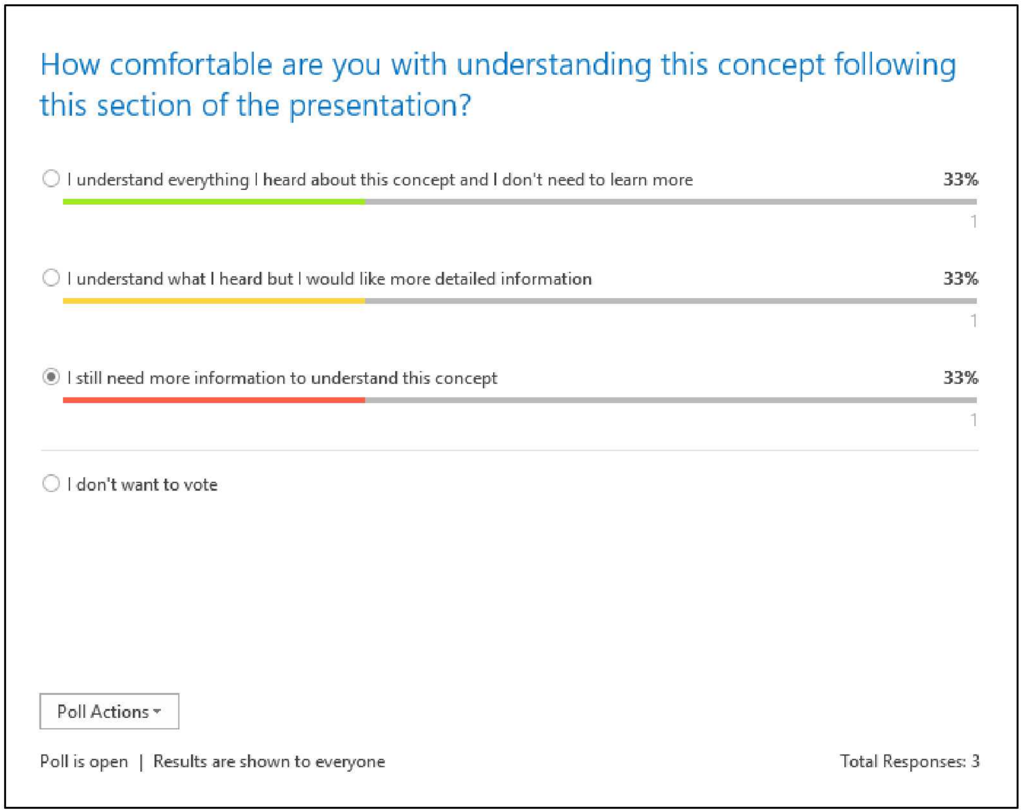
Physics and Material Model Fidelity Sub-elements

- Assess interpolation vs. extrapolation of physics and material model
  - To what extent does the application domain intersect the validation domain for this physics and material model (does not intersect, partially intersects, entirely contained)?



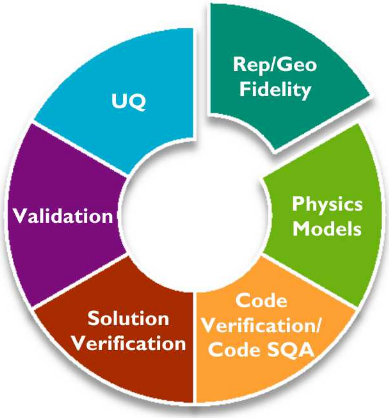
Real-time Feedback

- Concepts:
  - Meaning of physics and material model fidelity
  - What a PIRT is and how it is used



Representation and Geometric Fidelity

Representation and Geometric Fidelity



Representation and Geometric Fidelity – "How are geometric feature simplifications influencing simulation results?"

- The process of characterizing representation and geometric fidelity, identifying key simplifications, and assessing sensitivities.

As-Modeled

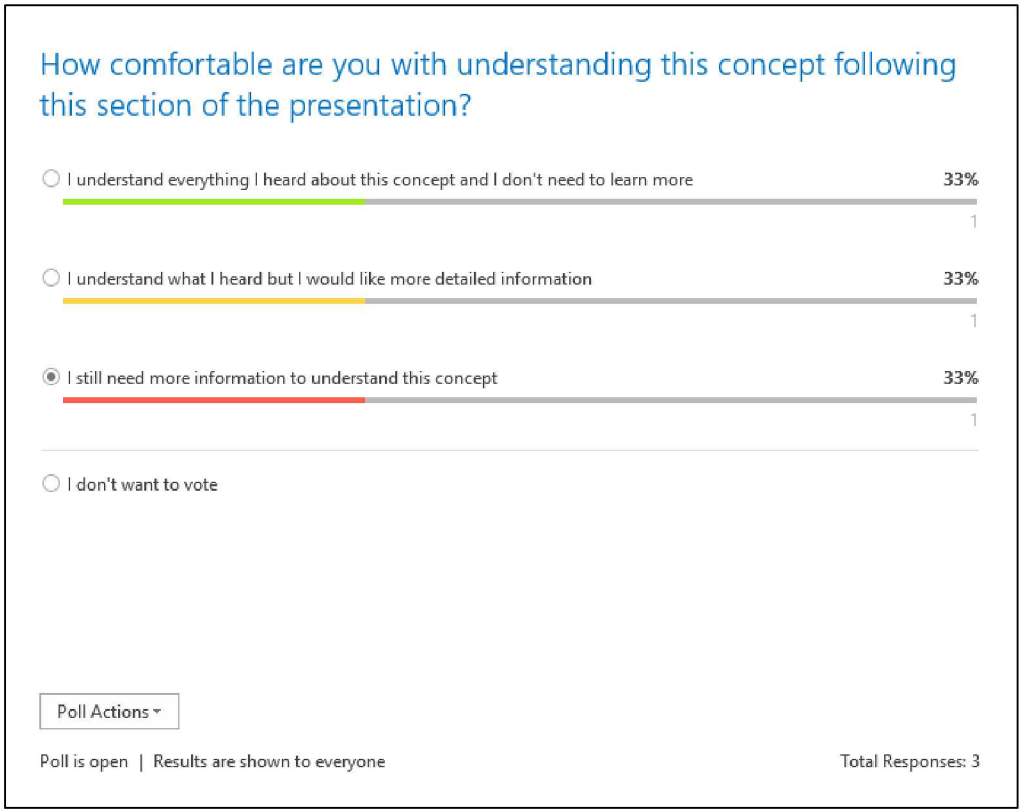
As-Designed

Representation and Geometric Fidelity Sub-elements

- Characterize Representation and Geometric Fidelity
  - Has the model been de-featured and to what extent are the "major" or "minor" features included (ex. Fillets, bolts, holes, cables, etc.)?
- Geometric Sensitivity
  - How is the computational error due to the given level of geometric resolution expected to influence the QOIs (perform simulations for varying levels of de-featuring)?
  - For which major features has the sensitivity been quantified (few, some, all)?

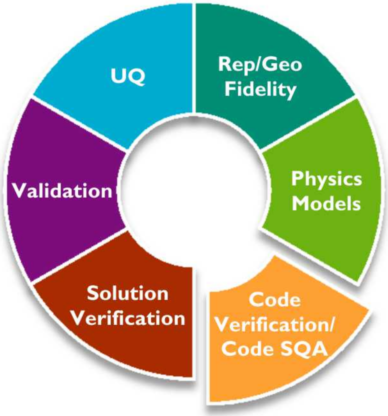
Real-time Feedback

- Concepts:
  - Meaning of representation and geometric fidelity
  - Importance of related credibility evidence



Code Verification

Code Verification

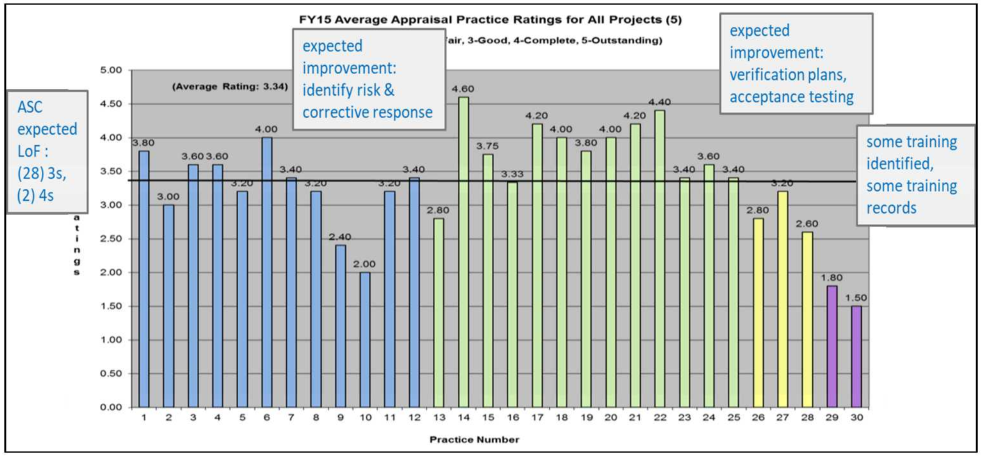


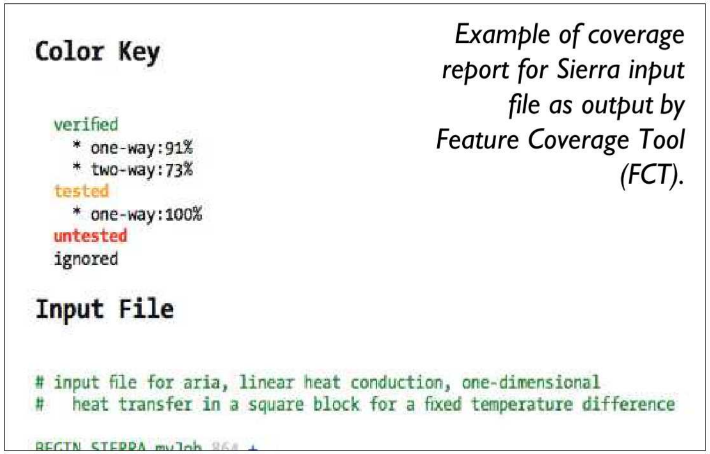
Code Verification - "Are there bugs in the code?"

The process of ensuring that the code correctly implements the numerical model.

- Errors in computer models are called code defects or bugs
- The code developers/testers have primary responsibility for identifying and eliminating code bugs

➤ Tool: Feature Coverage Tool (FCT)





Color Key

verified
* one-way:91%
* two-way:73%
tested
* one-way:100%
untested
ignored

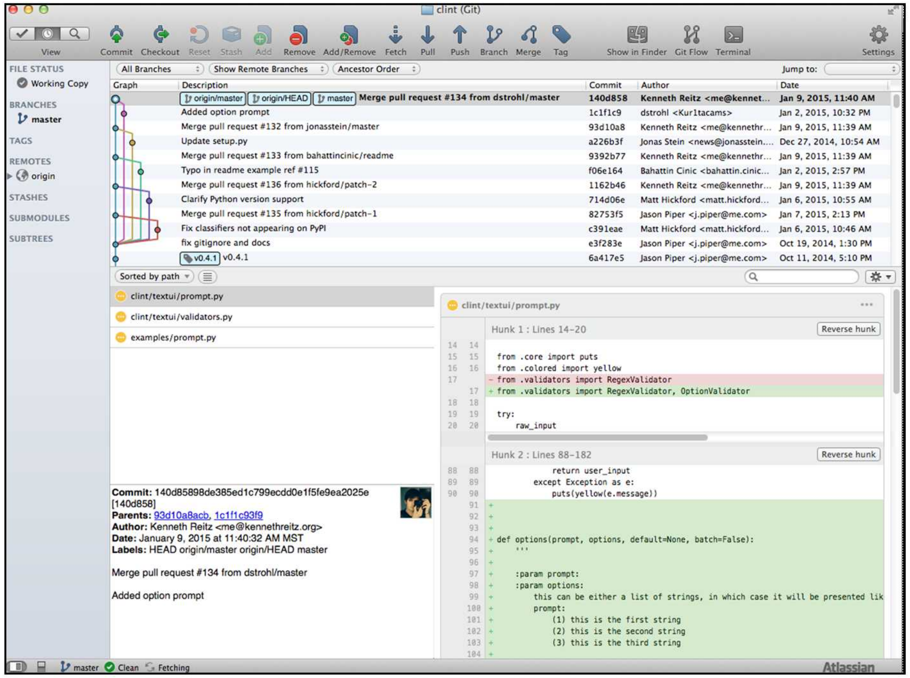
Input File

```
# input file for aria, linear heat conduction, one-dimensional
#   heat transfer in a square block for a fixed temperature difference
```

Example of coverage report for Sierra input file as output by Feature Coverage Tool (FCT).

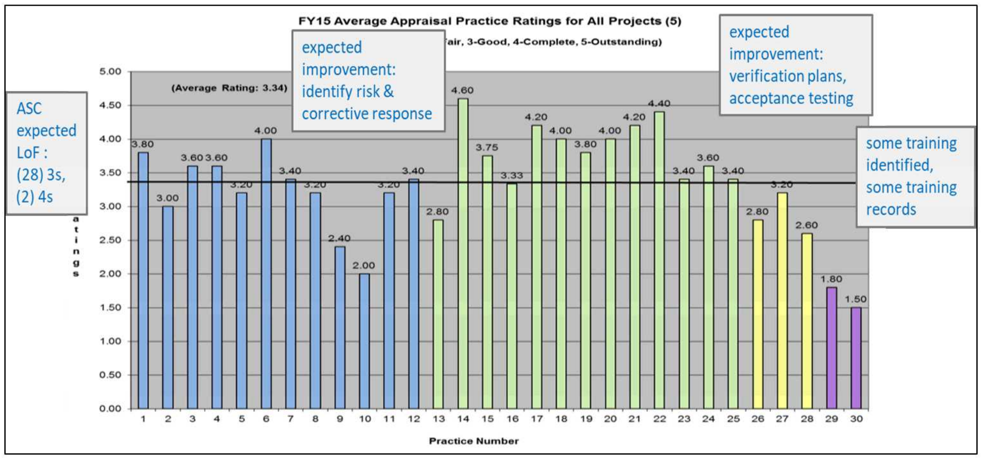
Code Verification Sub-elements

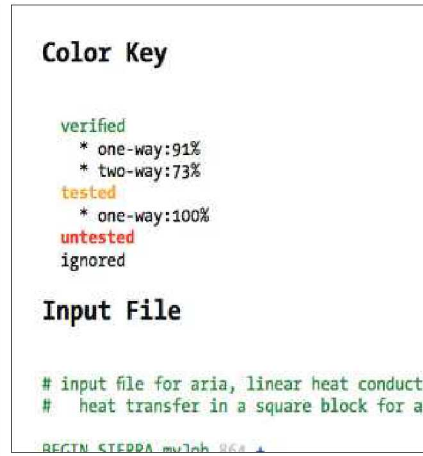
- Apply software quality engineering (SQE) processes (requires input from a capability developer)
  - Is the code capability managed to identified SQE practices? If so, reference them.
  - Is the SQE process managed and optimized?



Code Verification Sub-elements

- Provide test coverage information
  - What regression tests and verification test suite (VERTS) are available for the code capabilities?
  - How well are the code features required for the intended application covered by the VERTS?



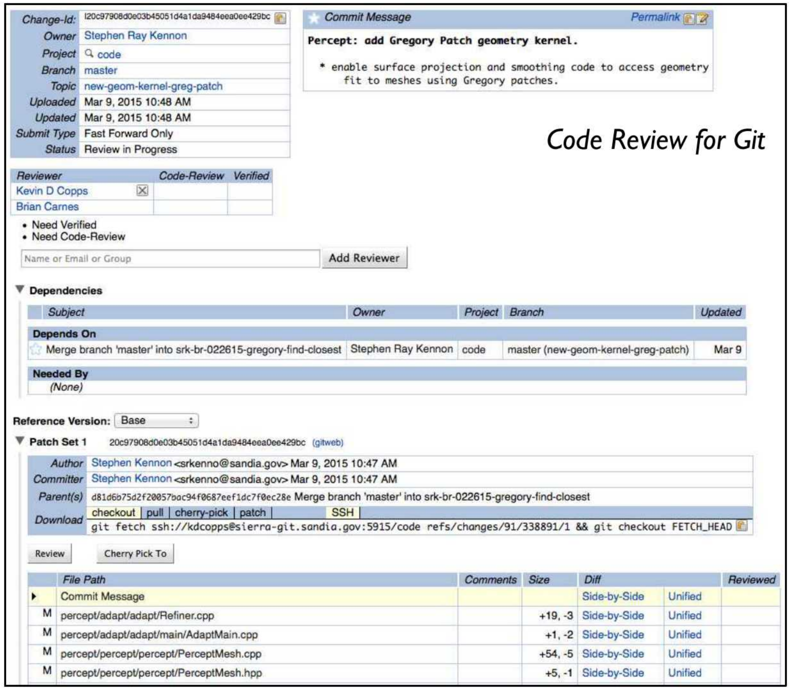


Example of coverage report for Sierra input file as output by Feature Coverage Tool (FCT).

- Identification of code or algorithm attributes, deficiencies and errors
  - How well are the code/algorithm attributes, deficiencies and errors from VERTS known?
  - How are these errors mapped to the intended application?

Code Verification Sub-elements

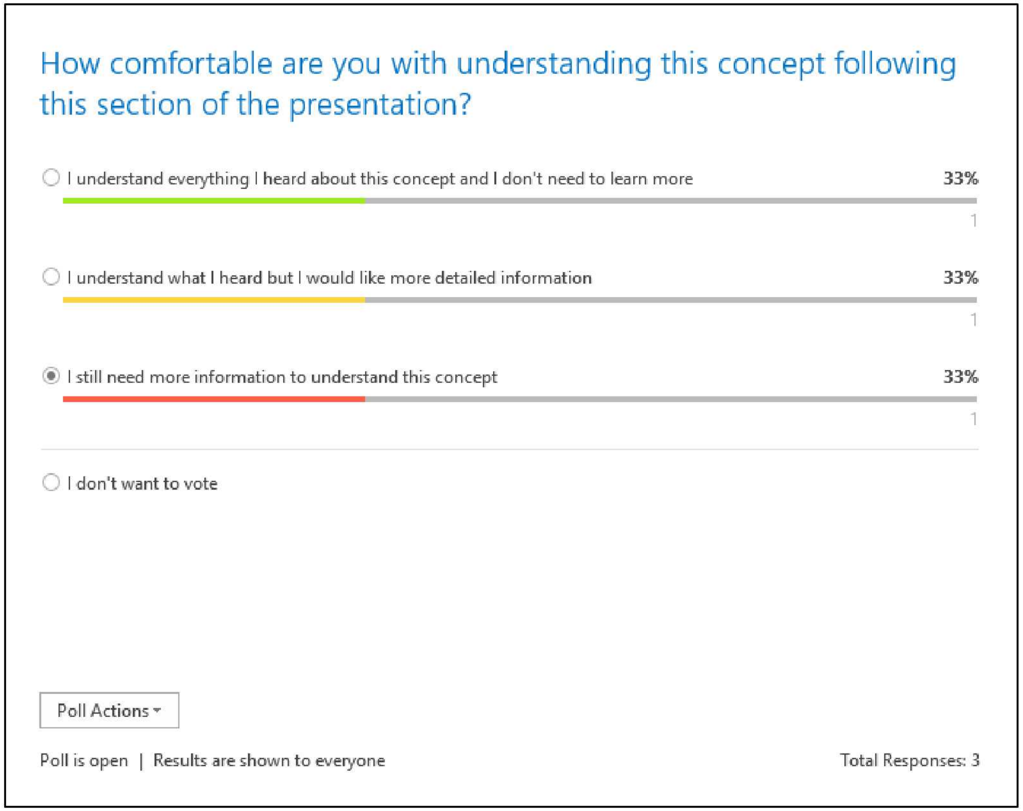
- Verify compliance to Software Quality Engineering (SQE) processes
  - How has the SQE process been reviewed (none, self-assessment, external, certification)?



Code Review for Git

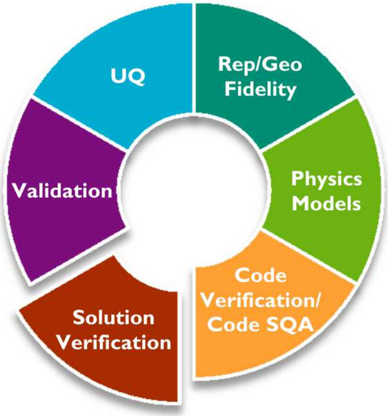
Real-time Feedback

- Concepts:
  - Code verification credibility activities
  - Importance of connecting code verification to the application and prediction of interest



Solution Verification

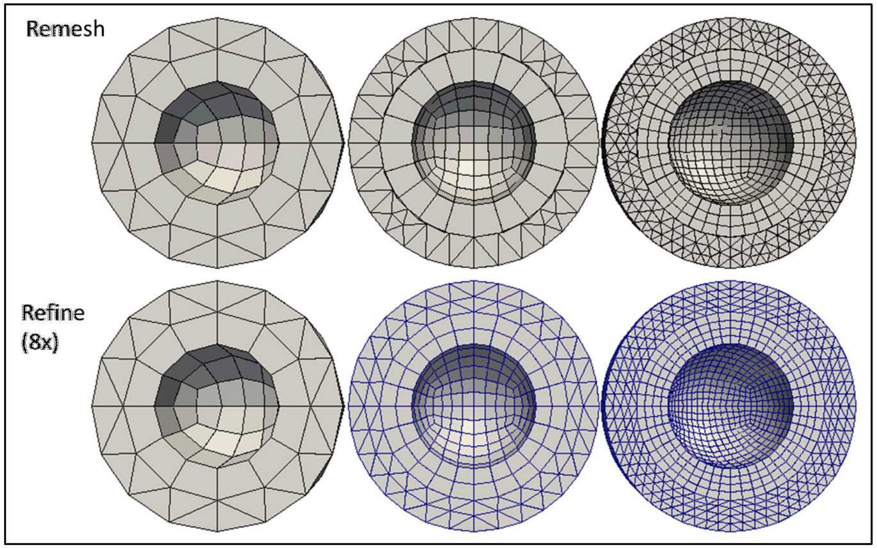
Solution Verification

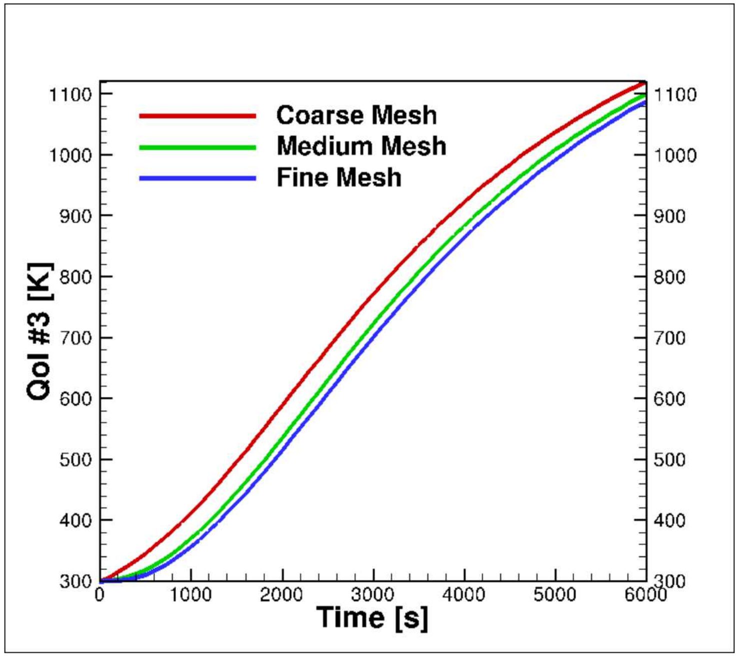


Solution Verification - "What is the numerical error?"

- The process of quantifying the numerical error in the computational simulation due to spatial discretization, temporal discretization, stochastic resolution, and iterative convergence.
- Done in the context of the overall uncertainty budget.
- Error may or may not need to be reduced.

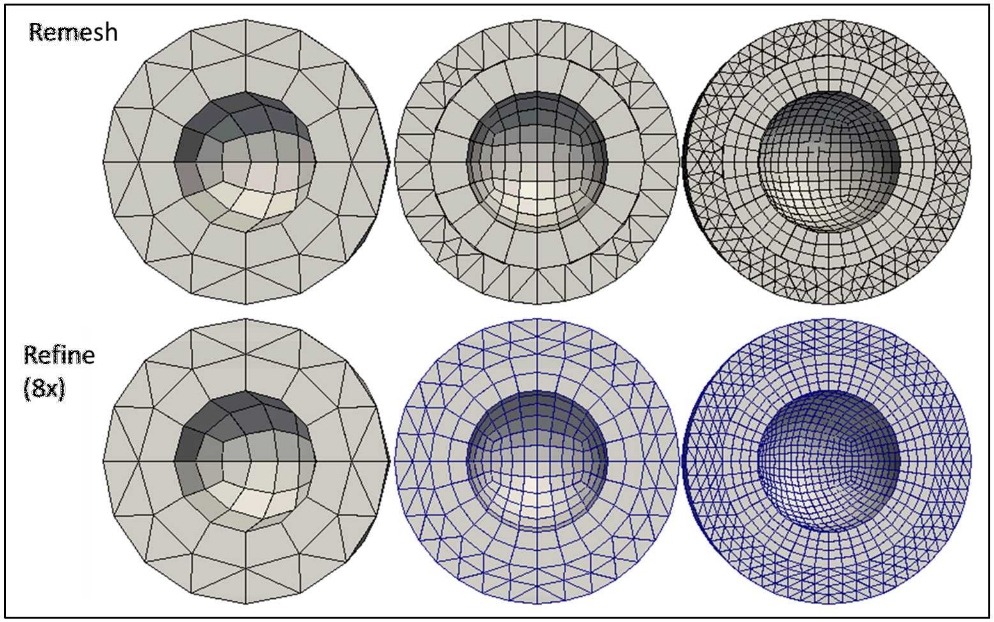
➢ Tool: Percept

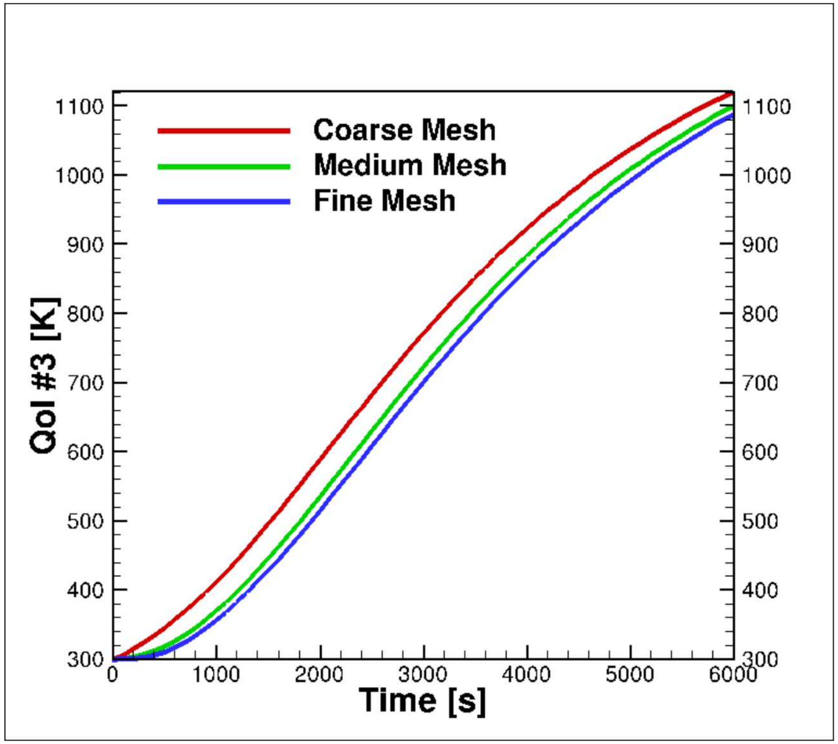




Solution Verification Sub-elements

- Quantify numerical solution errors
  - How have numerical errors incurred from spatial, temporal, and stochastic resolution been accounted for (qualitative vs. quantitative)?
  - How are these errors expected to impact all of the relevant QoIs?





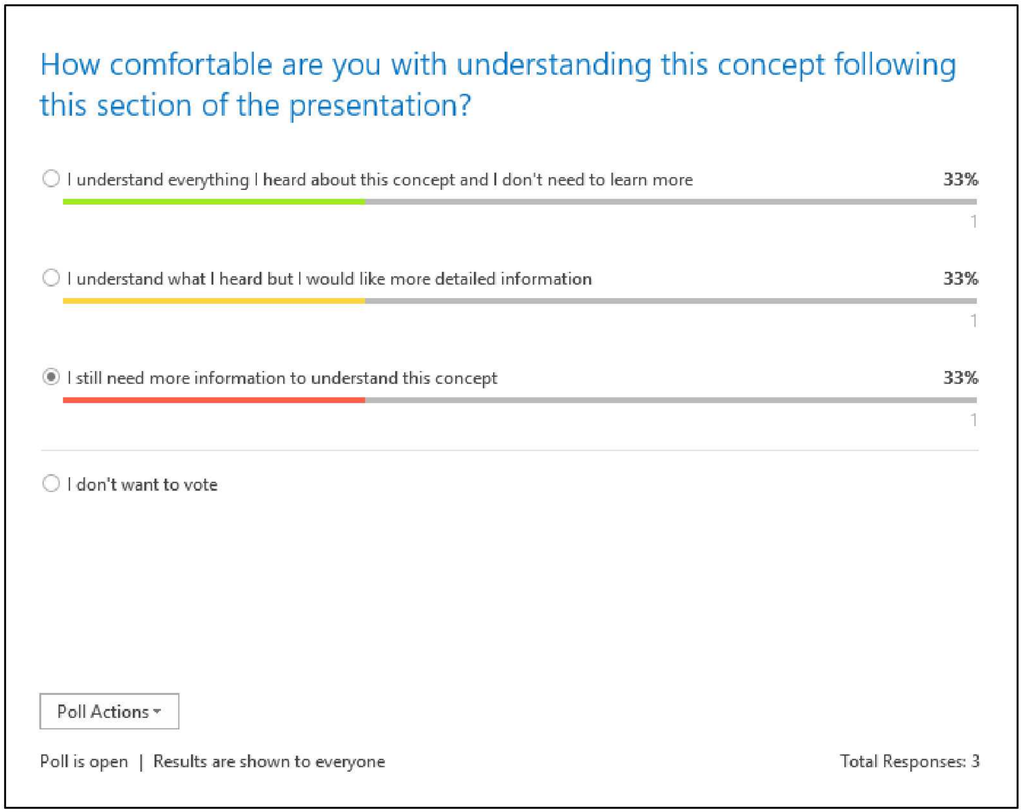
- Quantify uncertainty in computational (or numerical) error
  - Are there appropriate error bars for the stochastic error for all the relevant QoIs?

Solution Verification Sub-elements

- Verify simulation input decks
  - How and by whom has the accuracy of the input decks for the simulation been checked (by the analyst, by other analysts, by multiple other users)?
- Verify simulation post-processor input decks
  - Are a common set of post-processing tools used for the analysis, and are they held to a common set of SQE standards?
  - How and by whom has the accuracy of the inputs to the post-processing tools been checked (by the analyst, by other analysts, by multiple other users)?

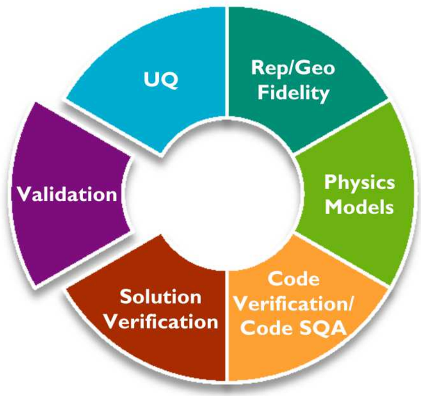
Real-time Feedback

- Concepts:
  - Solution verification credibility activities
  - Importance of performing solution verification prior to other credibility activities



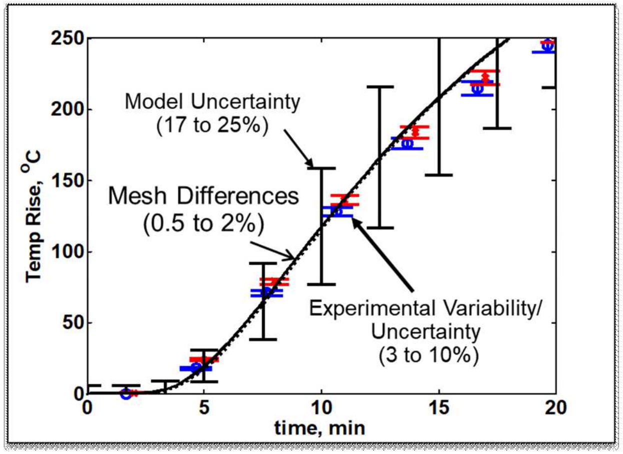
Validation

Validation



Validation – "Are we solving the right equations?"

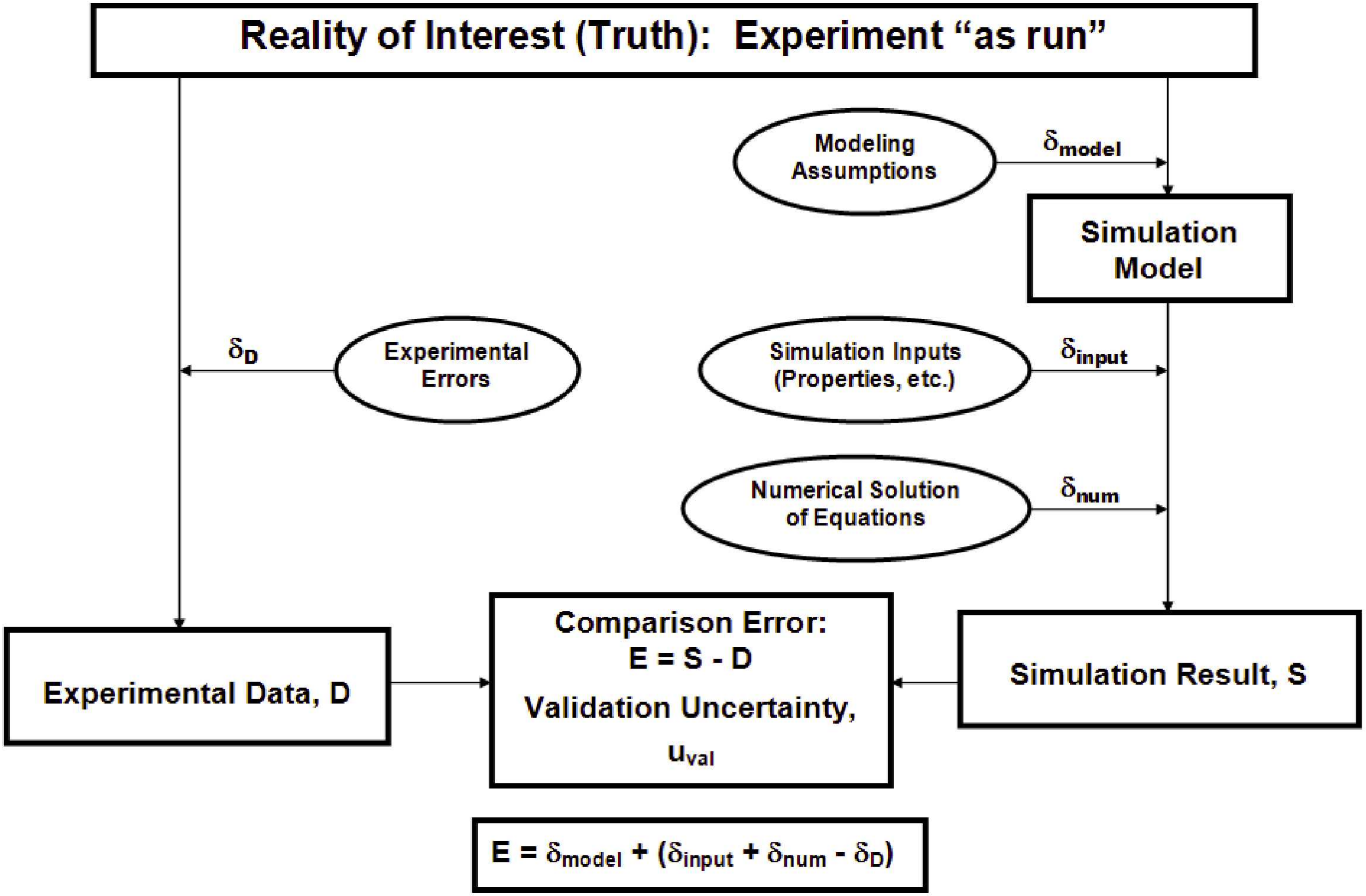
- The process of determining the accuracy of a computational simulation to represent the real world as approximated by experimental data.



- Model validation assesses a model in specific scenarios using experimental observation.
- Model validation quantifies the agreement between modeled prediction and truth <u>relative to the estimated uncertainty</u> of the validation exercise.

Validation Procedure

- Validation procedures have been described in standards and are used in practice across many applications



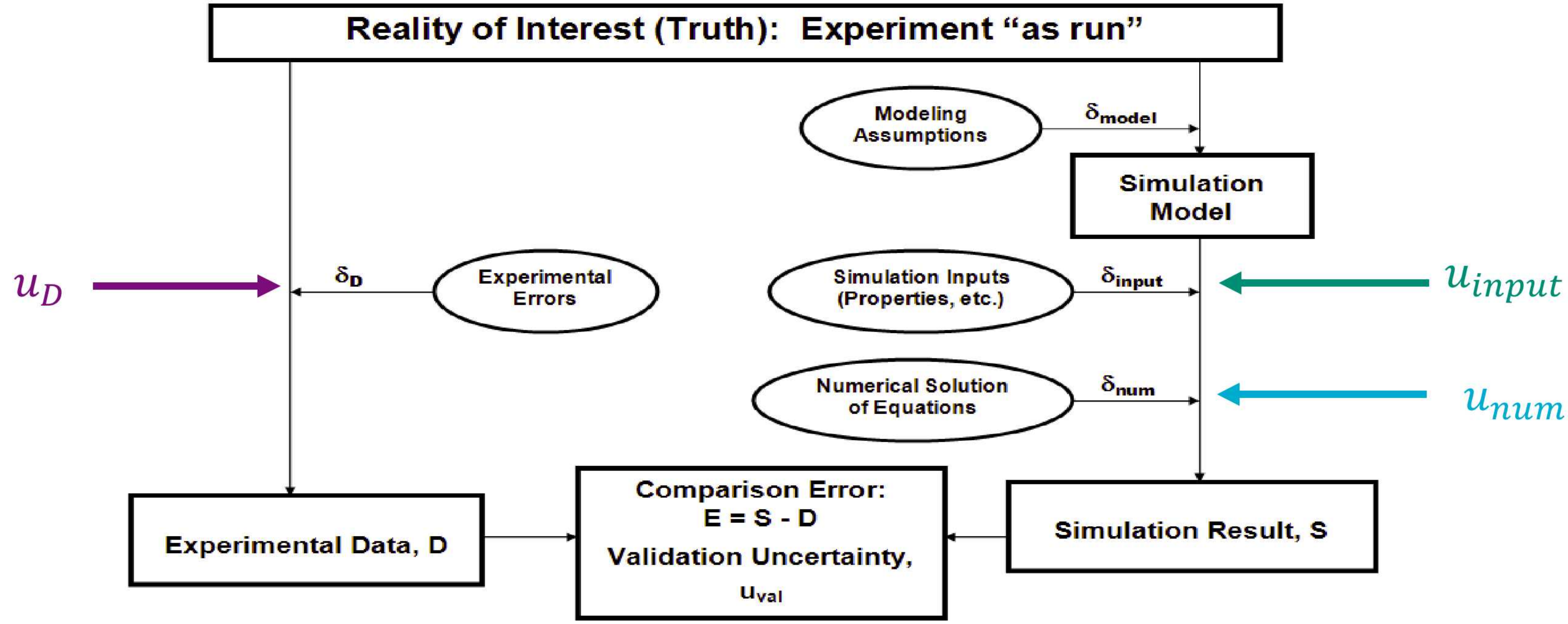
ASME V&V-20 (2009)

Validation Procedure

$$\delta_{\text{model}} \in \left[E - u_{val}, E + u_{val}\right]$$

$$E = S - D$$

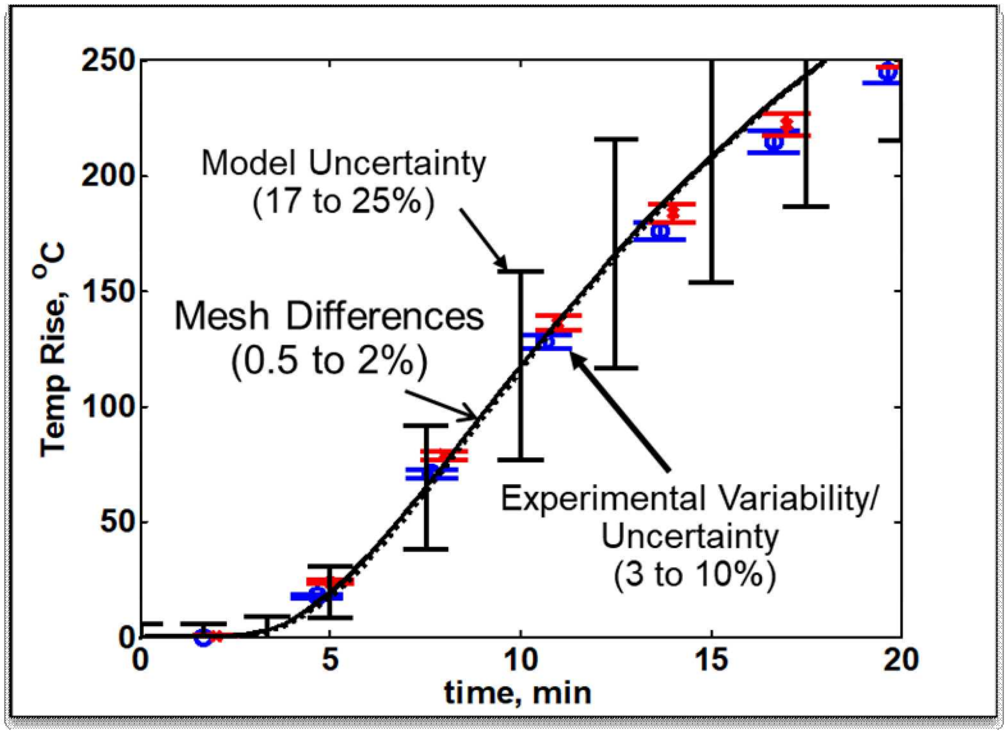
$$u_{val} = \sqrt{u_{num}^2 + u_{input}^2 + u_D^2}$$



ASME V&V-20 (2009)

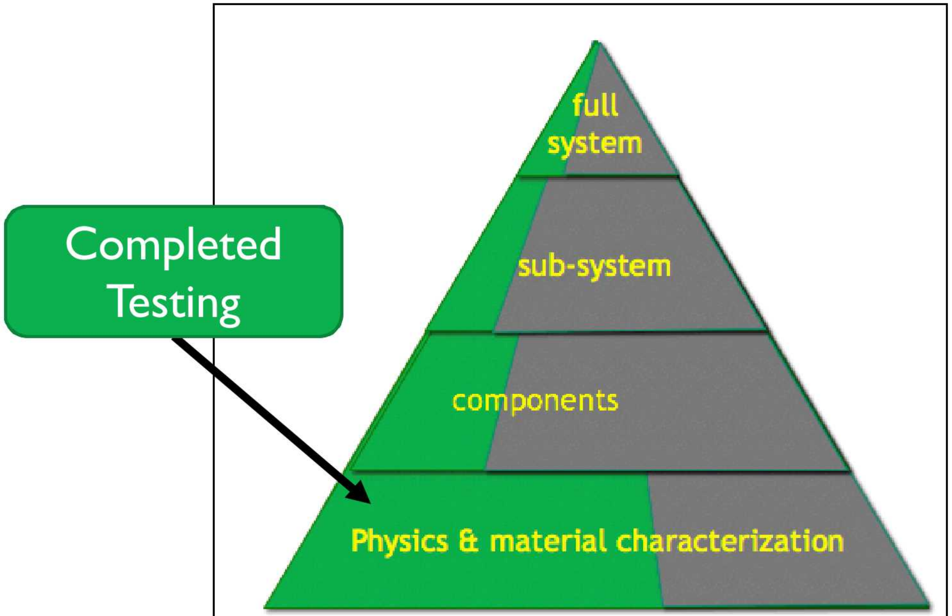
Validation Sub-elements

- Quantify physical accuracy
  - What is the rigor of the validation comparisons (i.e., are they quantitative or qualitative)?
  - Do the validation comparisons include uncertainty/error in the test data and model outputs?

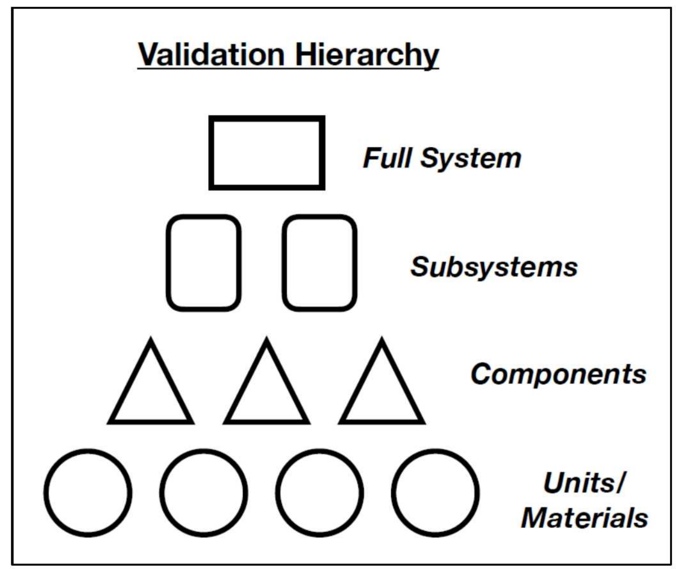


Validation Hierarchy

- A **Validation Hierarchy** maps from material to component to subsystem to full system levels and can be helpful for planning and execution of validation activities
- **Separate Effects** validation is used at the level of materials/basic physics to validate specific components of a model or code

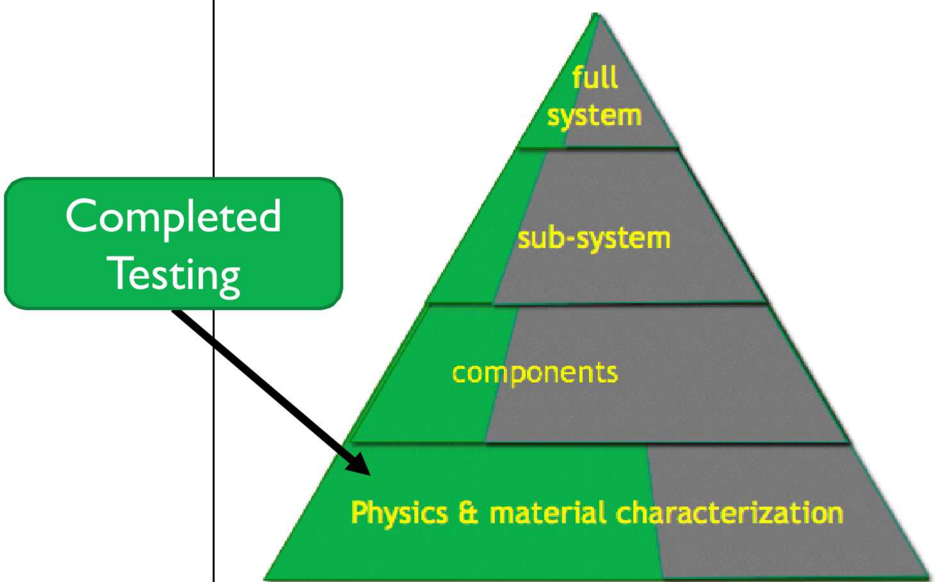


validation hierarchy

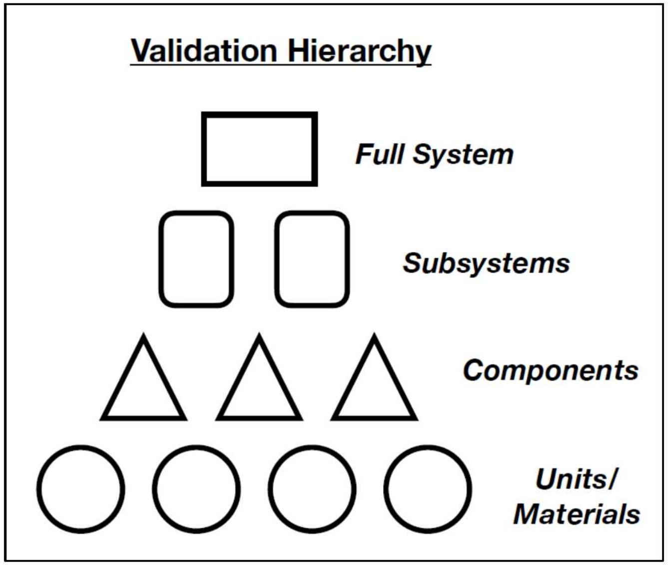


Validation Hierarchy

- Define a validation hierarchy
  - Has a validation hierarchy been defined (i.e., mapping from material to component to subsystem to full system levels)?
- Apply a validation hierarchy
  - What is the methodology for how available experimental data connects the levels of the hierarchy?
  - Have the steps in this methodology been performed (i.e., have quantitative comparisons been made at different levels of the hierarchy)?

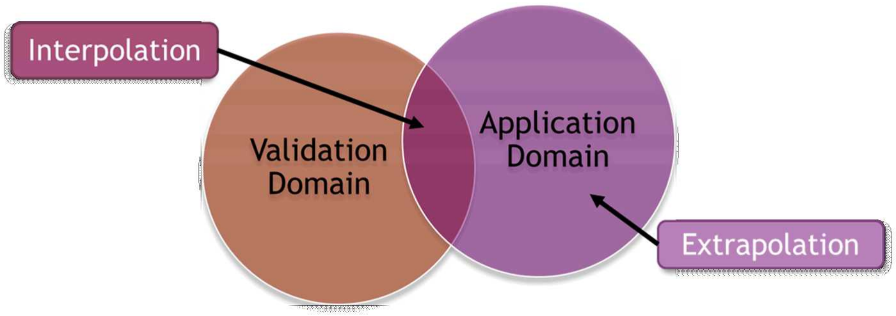


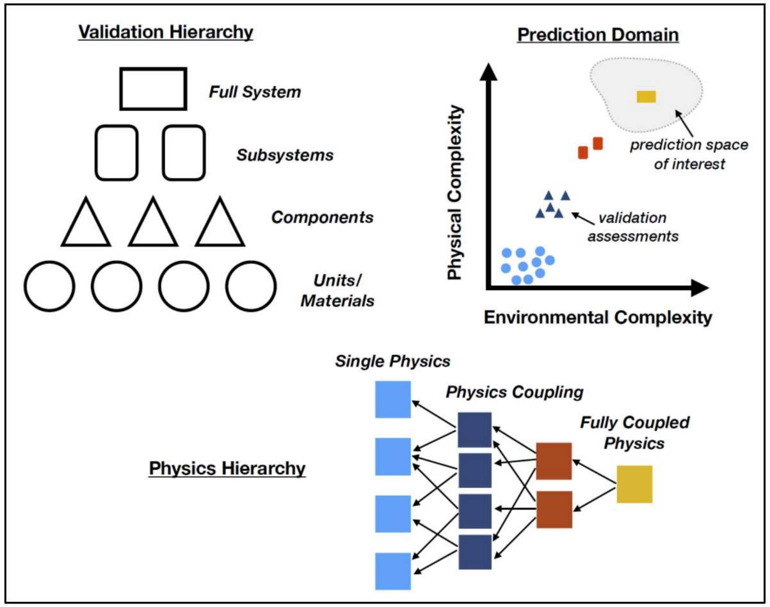
validation hierarchy



Validation Domain vs. Application Domain

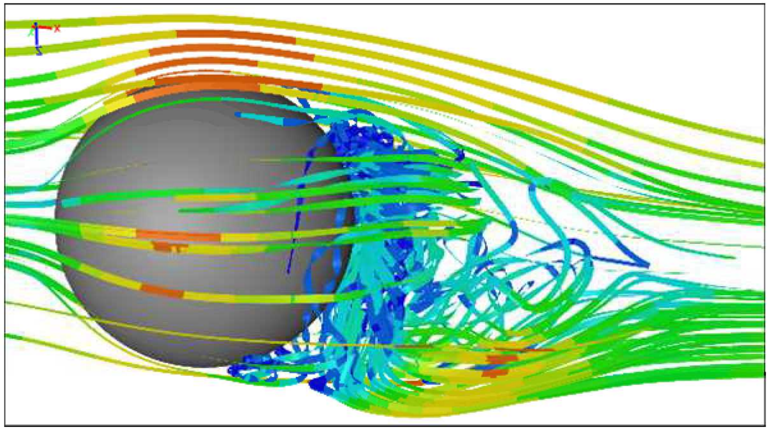
- Validation domain vs. application domain
  - Is the application of the model an extrapolation from the conditions where test data is available for validation, and to what extent (materials, environments, hardware, etc.)?
  - What evidence exists that provides confidence in the ability to extrapolate?





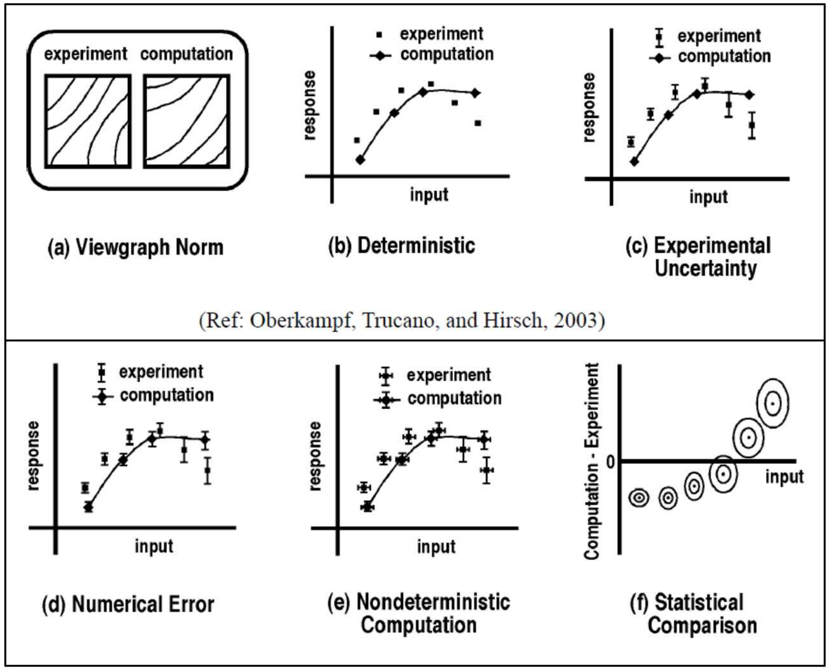
Experimental Credibility

- Experiments provide a **real-world view** of physics.
  - Come with cost and schedule.
- Simulations provide rapid insights at lower cost.
  - Come with potential model form and other errors.
- Experimentalists and modelers should work **collaboratively**, but good working relationships don't always happen naturally.
- The Experimental Credibility process facilitates discussion to **align goals** and **streamline efforts** between experimentalists and analysts while also guiding the collection of **experimental credibility** evidence.



Connections Between Analysis and Experiments

- Data are used in CompSim to:
  - Develop model forms
  - Calibrate model parameters
  - Validate model predictions
- Simulations **inherit the quality/credibility** of the experiments, including any errors present.
- The Experimental Credibility process provides a structured method to assess experiments used for simulations and includes:
  - Correctness
  - Completeness
  - Applicability to intended use
- It encourages:
  - Early planning of experiments
  - Communication between stakeholders: experimentalists, modelers, system integrator
  - Documentation of experimental credibility that aids simulation credibility



Validation comparison levels of rigor

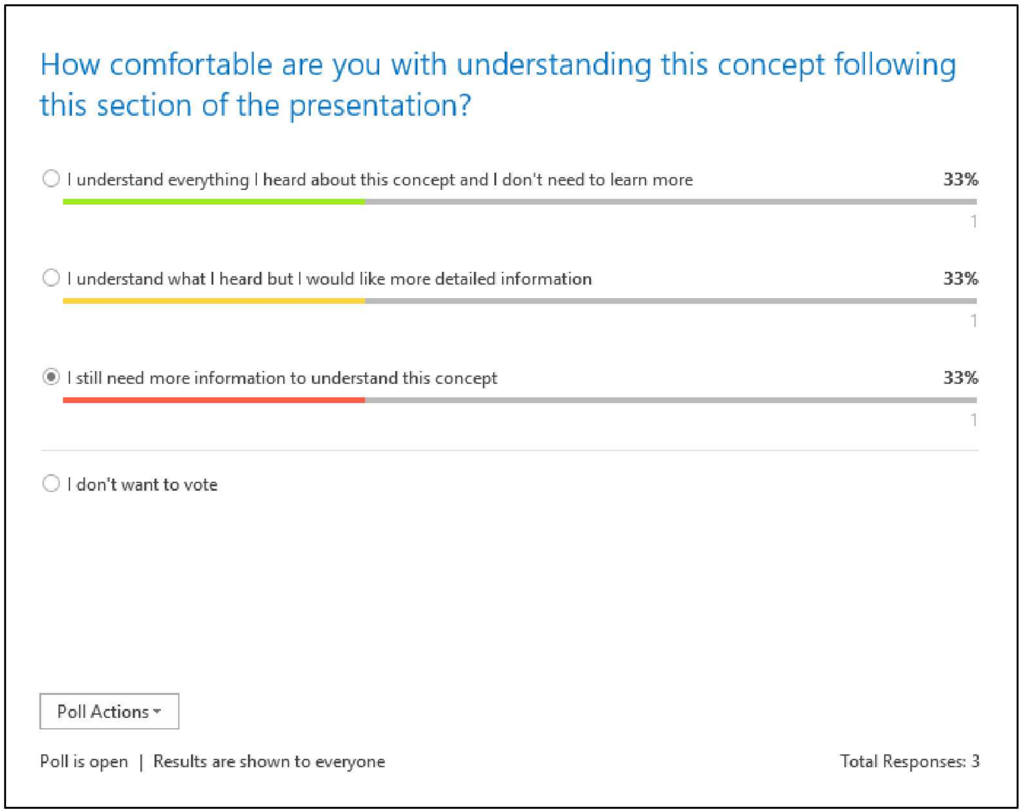
Plan and Assess Experiment Overview

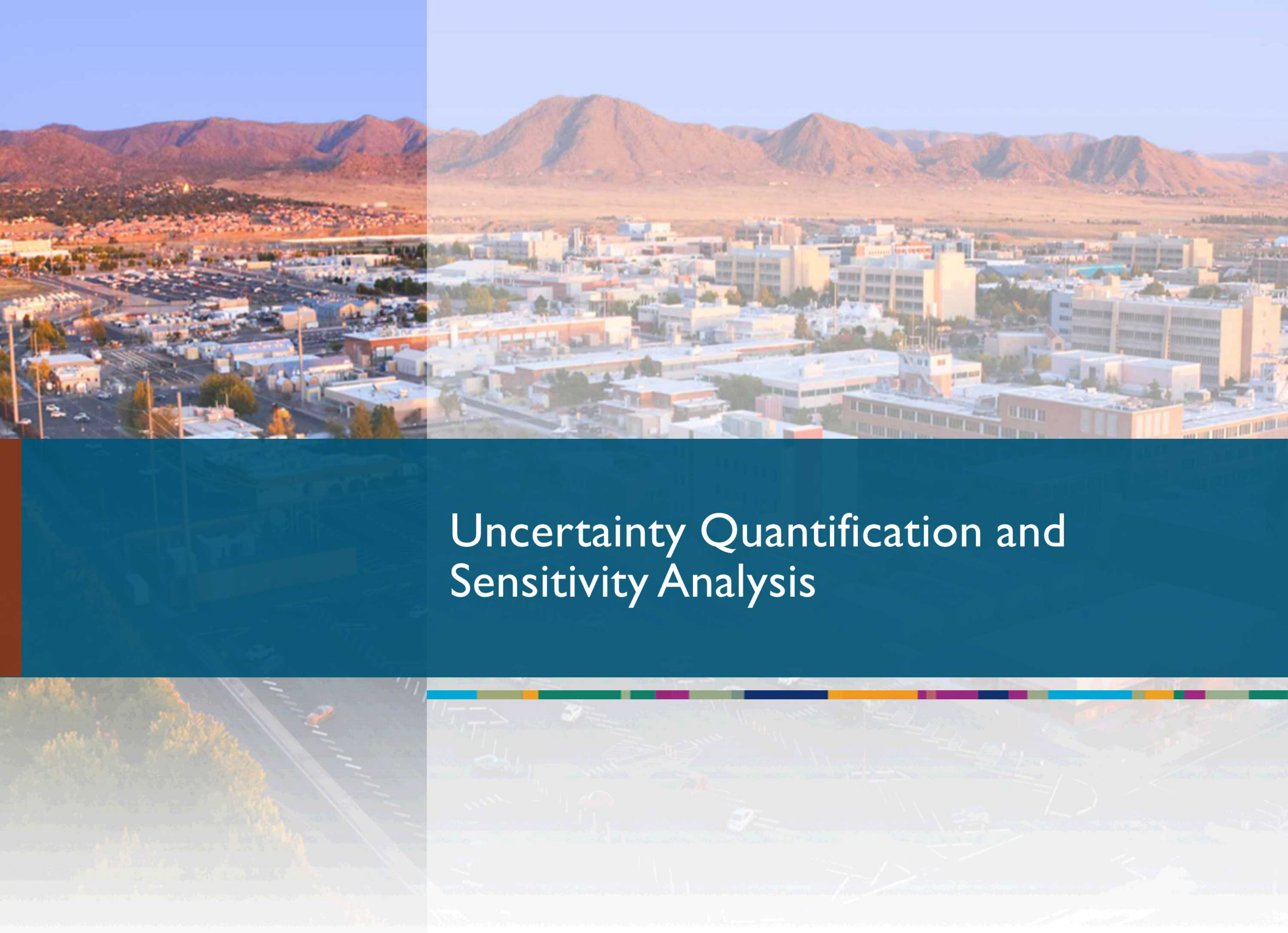
- Spreadsheet with seven elements
- Usable with any application
- The prompts are open-ended questions that contain best practices
- Team of experts and users
  - Modelers
  - Experimentalists
  - Customer
  - V&V partner
- Team assessment steps
  - Discuss prompts, strengths/weaknesses
  - Writes assessment commentary
  - Identify action items
- Elements
  - Planning
  - Intended Use
  - Sample, Geometric, Material Fidelity
  - Experimental & Environmental Fidelity
  - Experimental Verification
  - Uncertainty Quantification
  - Peer Review and Documentation

	Plan and/or Assess Experiment	Use when planning and/or assessing an experiment whose results are intended to inform computational simulation (CompSim) such as with calibration or validation activities. Read the prompts, discuss with team, and write a response for each element.	How did these elements impact the strength and weakness of this experiment/test for the purpose of the CompSim intended use?	If gaps are identified and actionable, list the assignment and person responsible
	Element	Prompts to Consider	Assessment Commentary	Action Items
Pre-Test	Planning	• What is the intended use of the test? Is it known to the experimentalist and analyst?		
		• How much communication is planned between the experimentalist, analyst, and customer/manager during the pre- and post-test stages? Are any adjustments appropriate?		
		• How will the analyst be involved in experiment planning? Could CompSim guide test planning?		
	Intended Use (e.g. materials characterization, calibration, validation)	• To what degree will the test conditions be characterized/measured for the intended use?		
		• To what degree will the test outputs be characterized/measured for the intended use?		
		• How will the measurement types and locations support the intended use?		
		• For validation, could metrics and acceptance criteria be specified early? Is testing over a range of parameter sets feasible to reveal trends?		
Pre- and Post-Test	Test Article Fidelity	• How relevant is the test article (material sample, model) to the application?		
		• To what degree is the test article pedigree known and documented, including any pre-processing?		
		• Are test article as-built (not as-designed) measurements available?		
	Test Condition Fidelity	• How relevant are the test conditions to the application?		
		• What could be changed to improve the applicability?		
		• How complete are test condition measurements to define CompSim inputs?		
	Experimental Verification	• How are test control and data acquisition methods verified?		
		• How are data post-processing scripts or processes verified? Can processing of synthetic data be used to identify errors?		
		• To what degree are test facility and instrumentation documented and		
		• To what degree could instrumentation alter test conditions?		
	Uncertainty Quantification	• To what degree were experimental uncertainties quantified for boundary and initial conditions as well as test outputs?		
		• How could test repeatability and/or person-person variability be assessed?		
		• For more on uncertainty quantification, see the "Assess Experimental Uncertainty" spreadsheet.		
	Peer Review and Documentation	• Which of the other elements will be reviewed by a subject matter expert?		
		• Which of the other elements will be documented further and to what degree?		
		• Are tabulated test data linked to their description, archived, and accessible?		

Real-time Feedback

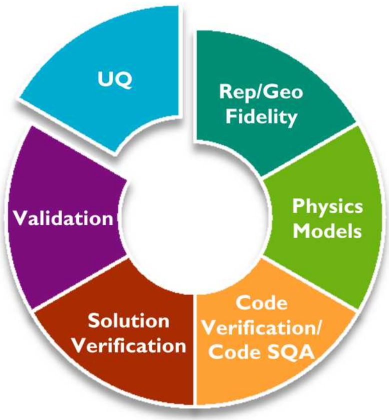
- Concepts:
  - Validation credibility activities
  - Concepts of a validation hierarchy and understanding the relationship between the validation and application domains





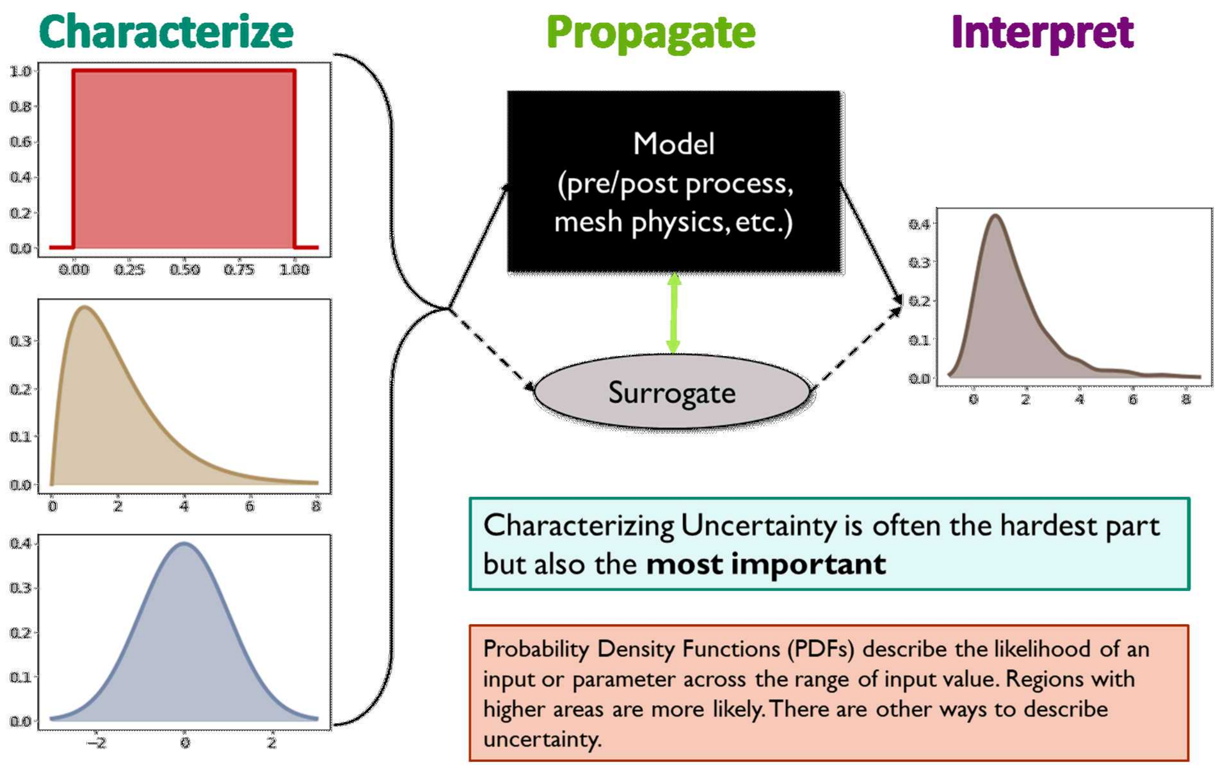
Uncertainty Quantification and Sensitivity Analysis

Uncertainty Quantification



Uncertainty Quantification (UQ) – "How large is the uncertainty in the result?"

- The process of characterizing all relevant uncertainties in a model and quantifying their effect on a quantity of interest.
- Sensitivity Analysis (SA) is an important component



Characterizing Uncertainty is often the hardest part but also the **most important**

Probability Density Functions (PDFs) describe the likelihood of an input or parameter across the range of input value. Regions with higher areas are more likely. There are other ways to describe uncertainty.

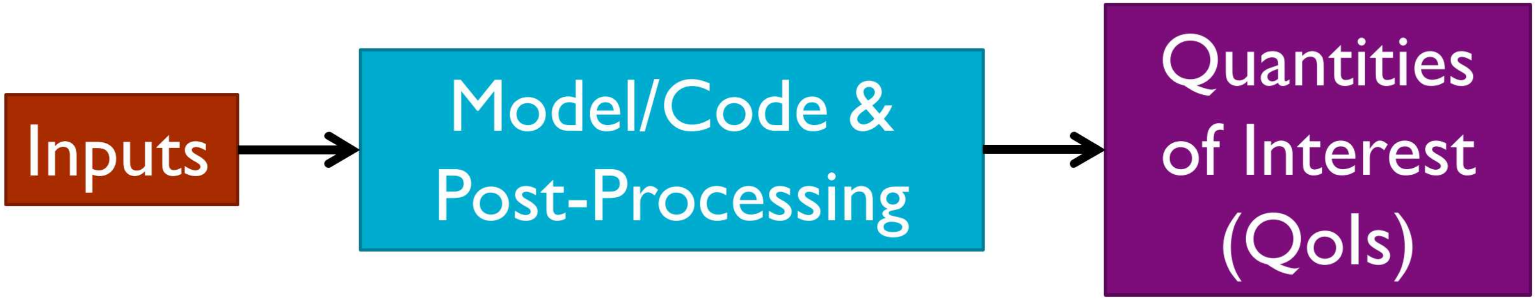
Error and Uncertainty

- **Error** is the result of a measurement (or simulation) minus a true value. The sign and magnitude of an error are typically unknown.
  - Error is the answer to the question: How far away is this measurement or simulation result from a true value?

- **Uncertainty** is a quantity associated with the result of a measurement (or simulation) that characterizes the dispersion of the values that could be attributed to the measurand (or simulation).
  - Uncertainty is a quantity to characterize error.
  - Uncertainty is the answer to the question: What is the range of values that we expect to encompass the true value of the system behavior?

Abstract View of Simulation

- Think of models / code as a black box
  - Model/Code and post-processing define a relationship between inputs and outputs



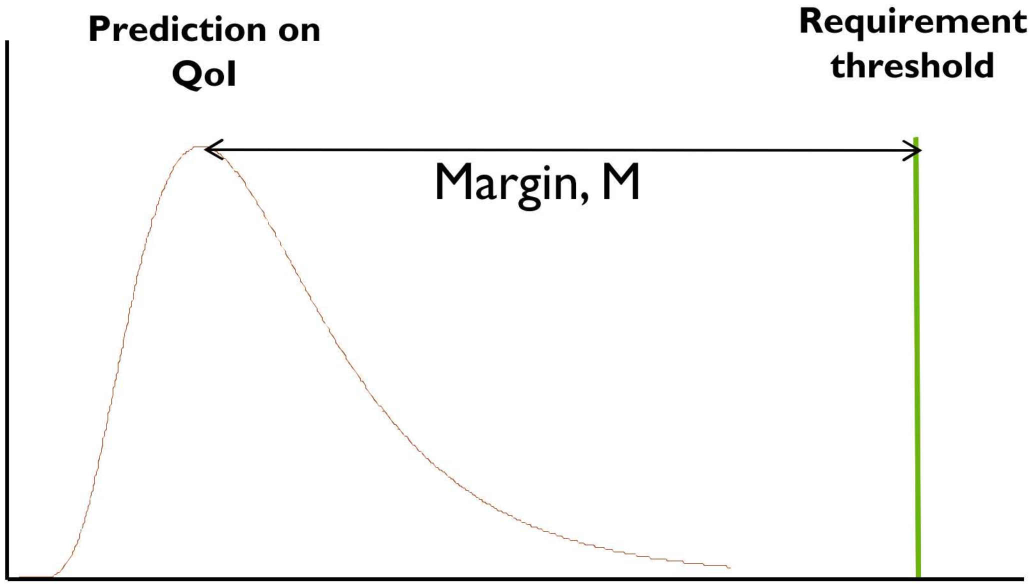
- Input = Anything that changes the QoIs
  - Model parameters, code settings (solvers, tolerances)
  - Boundary conditions, external forcing, etc.
  - Mesh and geometry
  - Computational hardware

Quantities of Interest

- A **QoI** is a quantity that is directly tied to a requirement or regulatory decision. It can be a model output or a function of a model output.
- Example QoIs resulting from UQ:
  - **Mean** Shielding Effectiveness (SE)
  - **99th percentile** of SE
- Example requirements based on probabilistic QoIs:
  - The **mean** SE should fall **below** XX dB
  - The **99th percentile** of SE should fall **below** XX dB with **95% confidence**
- For the purposes of UQ, the choice of QoI will affect:
  - Whether to separate aleatory and epistemic uncertainties
  - The number of realizations needed

QMU

- Some requirements are centered around **Q**uantifying **M**argins of performance for given conditions in the presence of **U**ncertainty (**QMU**)



Translation

Given that:
We don't know exactly how a system behaves and
We don't know exactly what the initial conditions, boundary conditions, and environment will be
Can we claim that the system will meet requirements?
By how much?
How confident are we? Are the results credible?

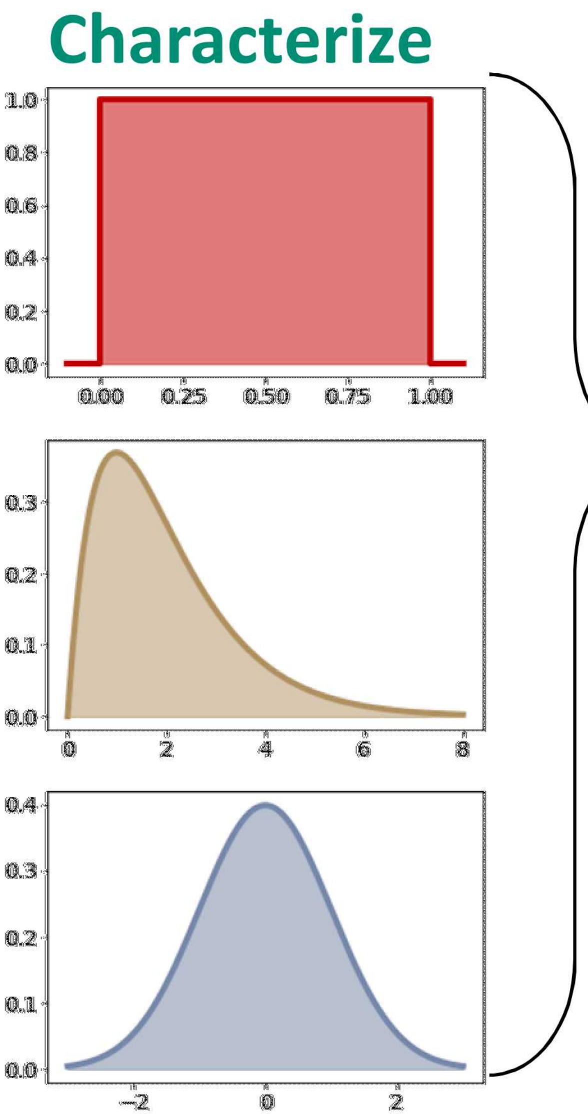
- A common challenge: Limited test data to characterize uncertainty and/or cannot test over the full application domain

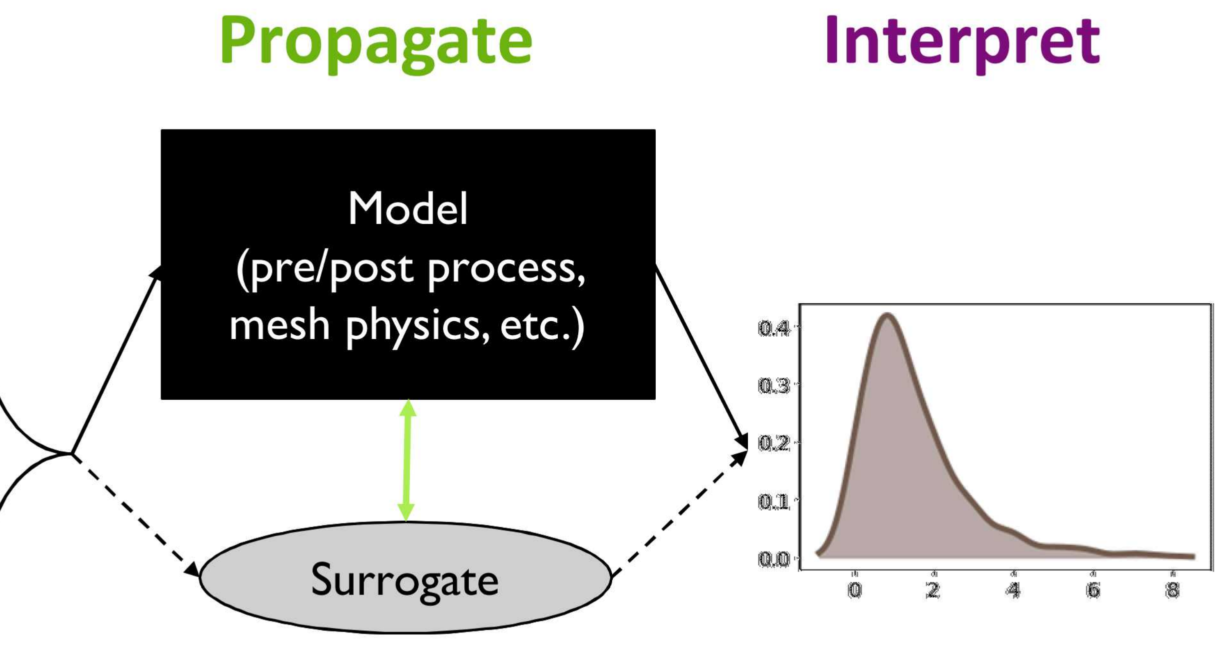
Uncertainty Quantification (UQ)

- The process of characterizing all relevant uncertainties in a model and of quantifying their affect on a QoI – ASME V&V 10, 2006
- What is uncertainty?
  - Lack of information or inherent randomness
- Uncertainty quantification = information quantification
  - Have a model, know the significant inputs, etc.
  - How much information do you have about QoI's?
  - What are the significant sources of uncertainty?
- Uncertainty quantification steps:

1. **Characterize the uncertainty for *significant* inputs**
2. **Propagate the uncertainty through the model**
3. **Interpret the resulting uncertainty**

Uncertainty Quantification





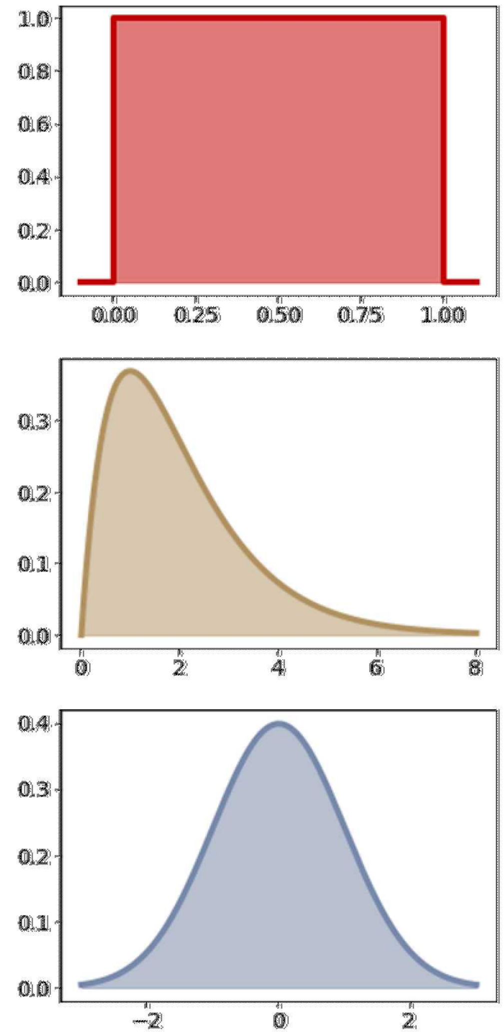
Characterizing Uncertainty is often the hardest part but also the **most important**

Probability Density Functions (PDFs) describe the likelihood of an input or parameter across the range of input value. Regions with higher areas are more likely. There are other ways to describe uncertainty.

Characterization of Uncertainty

- Sources of Uncertainty
  - Model parameters
  - Mesh, geometry
  - Experimental conditions (controlled/uncontrolled variables and boundary conditions)
  - Experimental data (measurement error/data sparseness)
  - Physics model form error (competing models)
  - Model parameters
  - Code errors, solver settings and solution approximations (numerical uncertainty)
- Types of uncertainty
  - **Epistemic** and **Aleatory**
  - Provide more insight into the information we have
- Quantitative methods
  - Must provide a mathematical description of parameters (and *input* uncertainty)

Characterize



Review of Aleatory and Epistemic Uncertainties

Epistemic (Reducible) Uncertainty

- Due to a lack of knowledge about appropriate value to use
- Can be reduced through increased understanding or more data
- Examples: Insufficient experimental data to characterize a probability distribution, poor understanding of physics

Aleatory (Irreducible) Uncertainty

- Random variability that cannot be reduced through further knowledge/data
- Examples: Part-to-part variation, weather variability

A parameter can have both types of uncertainty

Propagation of Uncertainty

- Intrusive/Embedded:
  - Represent the parameters stochastically
  - Rewrite ***all*** code to properly handle stochastic parameters
  - Requires only a single, but **much** more expensive solution

Possibly not practical for all but the simplest models

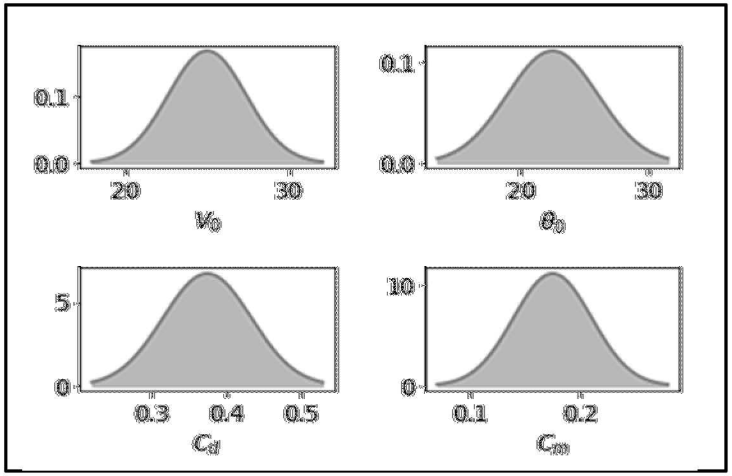
- Non-Intrusive:
  - Sampling methods (e.g., Monte Carlo (SRS, LHS))
    - Evaluate response from ensemble of samples
  - Polynomial Chaos Expansion (PCE)
  - MOST other methods can be formulated as
    1. Construct a **surrogate model** with as few realizations as possible
    2. Sample the surrogate model with many, many more samples
    3. Compute desired quantities
  - **Surrogate Models** (also known as emulators, meta-models, and response surfaces) are relatively fast statistical models that approximate more complex computer models.

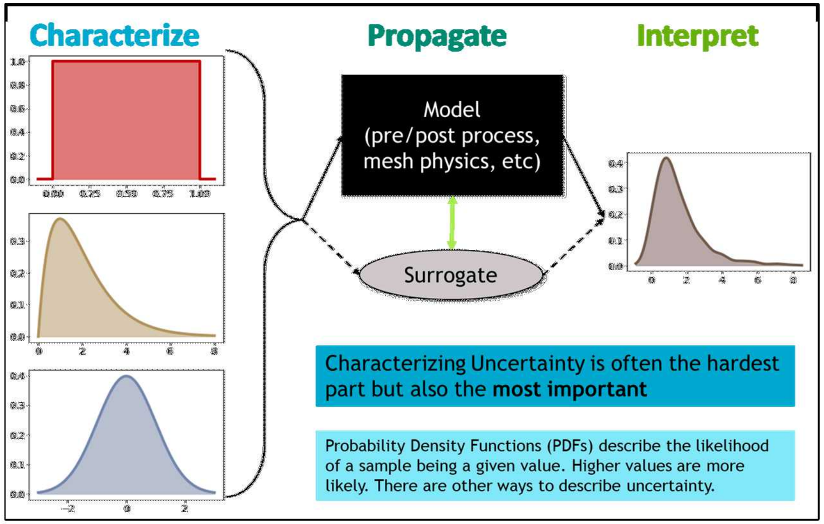
Treat existing models as a "black box" and propagate uncertainty by evaluating the model for different settings

DAKOTA

Uncertainty Quantification Sub-elements

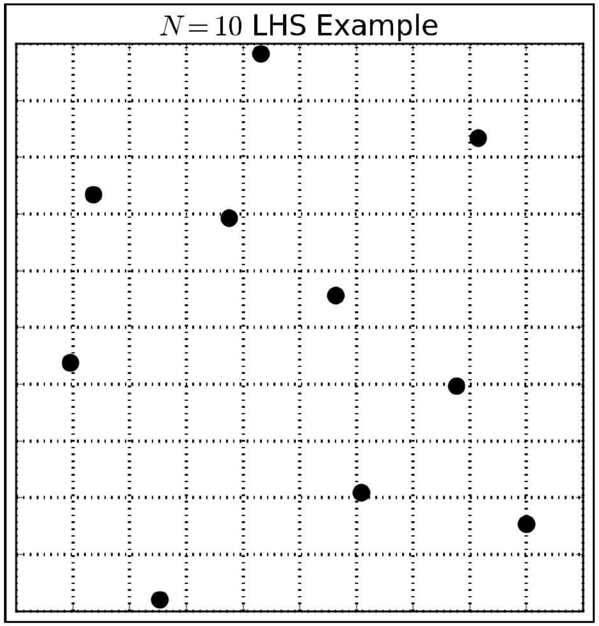
- Aleatory and epistemic uncertainties identified and characterized
  - Aleatory = natural variability; epistemic = lack of knowledge
  - Has an inventory of uncertainty sources been taken, and have they been classified according to these forms?
  - What is the source of information (e.g., legacy, literature, direct measurement, calibration, etc.) that is used for uncertainty characterization (e.g., classification as aleatory vs. epistemic, uncertainty representation, distributional assumptions, etc.)?

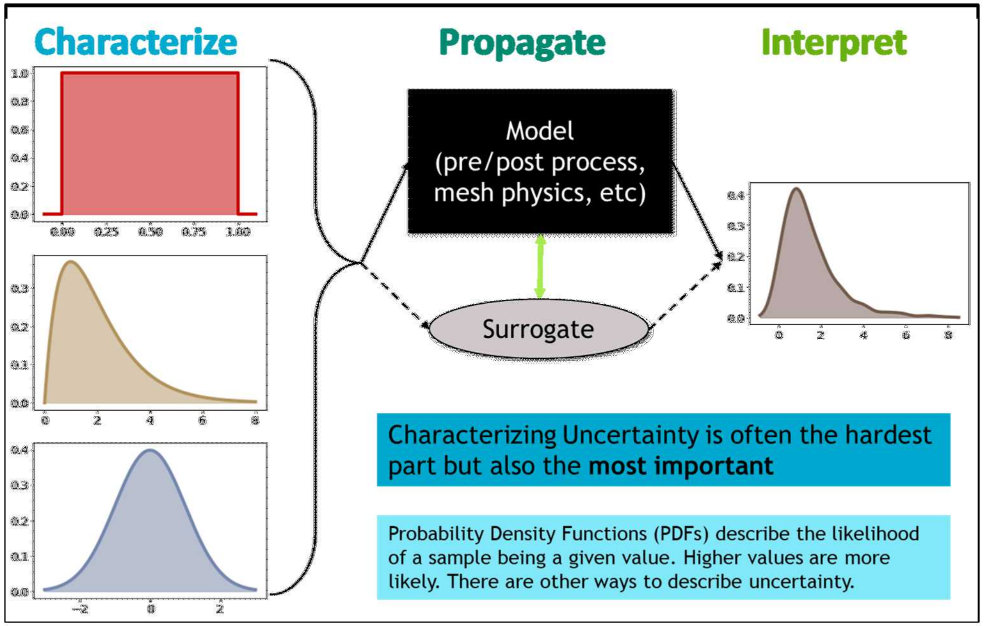




Uncertainty Quantification Sub-elements

- Quantify impact of uncertainties on QoIs
  - Have identified sources of uncertainty (see 1 above) been propagated to the important output QoIs?
  - What is the procedure for propagation and what additional errors are introduced?

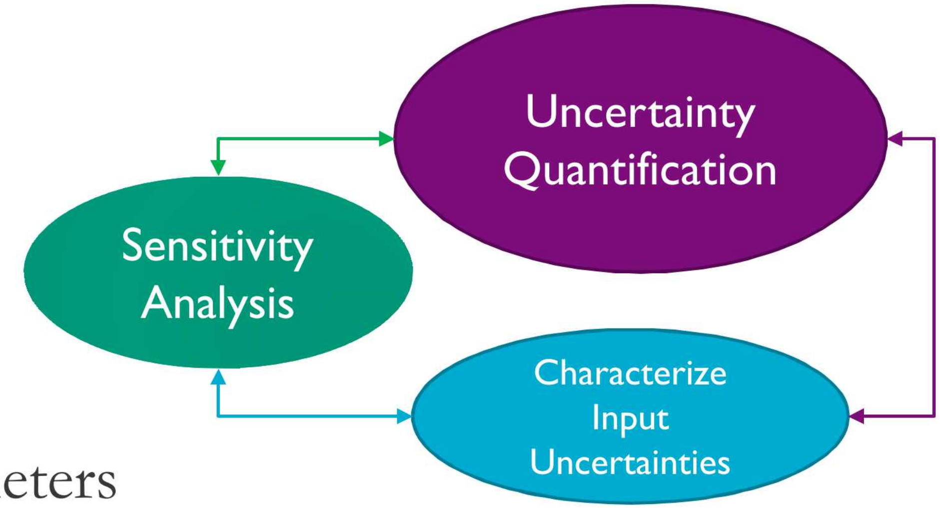




Sensitivity Analysis (SA)

How do changes to inputs affect the response?

- How "sensitive" is the response to each input?
- Direction and magnitude
- Which inputs matter the most?
  - Can we ignore any sources of uncertainty?



Typically focus on model parameters OR other inputs

Sensitivity Analysis and Uncertainty Quantification are *iterative*

Local vs. Global Sensitivity Analysis

- **Local** sensitivity analysis
  - Compare local relative derivatives – these can change over a domain of interest
  - Most sensitive ≠ most significant
    - Should consider actual range of parameters (if known)
  - Number of model evaluations is usually less than global methods

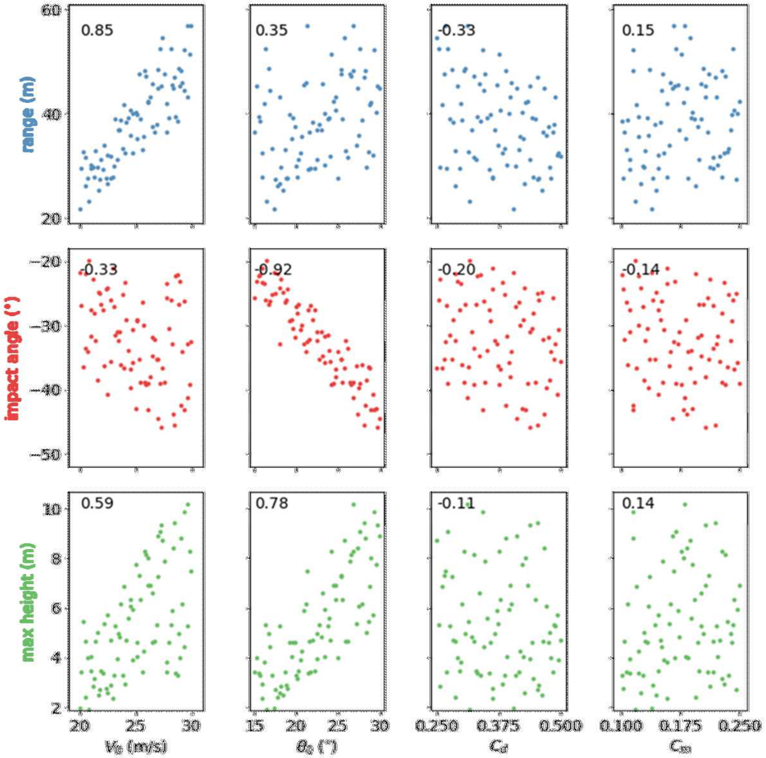
- **Global** sensitivity analysis
  - Consider cumulative affect across the domain
  - Correlation-based (correlation coefficients) or variance-based (Sobol indices)
  - Sampling provides a more global picture of model response
  - Results may depend on parameter characterizations (ranges, etc.)

Uncertainty Quantification Sub-elements

- Perform sensitivity analysis
  - How have the most important uncertainty sources for the relevant QoIs been identified (e.g., SME judgment, local sensitivity analysis, global sensitivity analysis, etc.)?

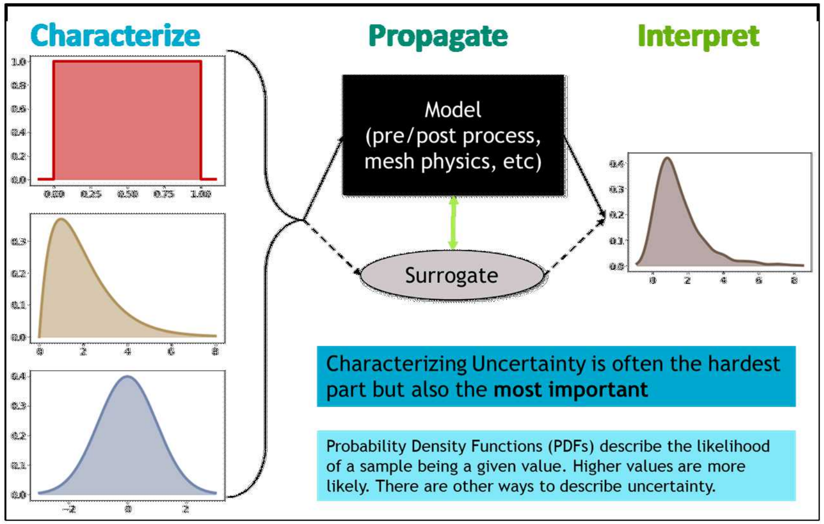
Sensitivity analysis using correlation coefficients for projectile problem.

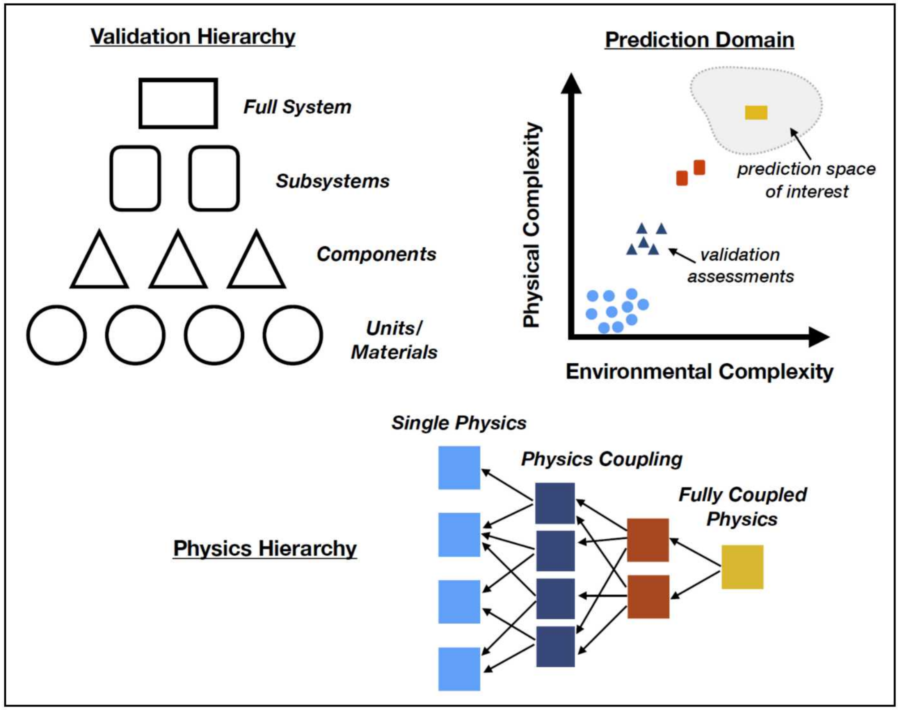
	Range (m)	Impact angle (°)	max height (m)
V_0 (m/s)	+0.85	-0.33	+0.59
θ_0	+0.35	-0.92	+0.78
C_d	-0.33	-0.20	-0.11
C_m	+0.15	0.14	+0.14



Uncertainty Quantification Sub-elements

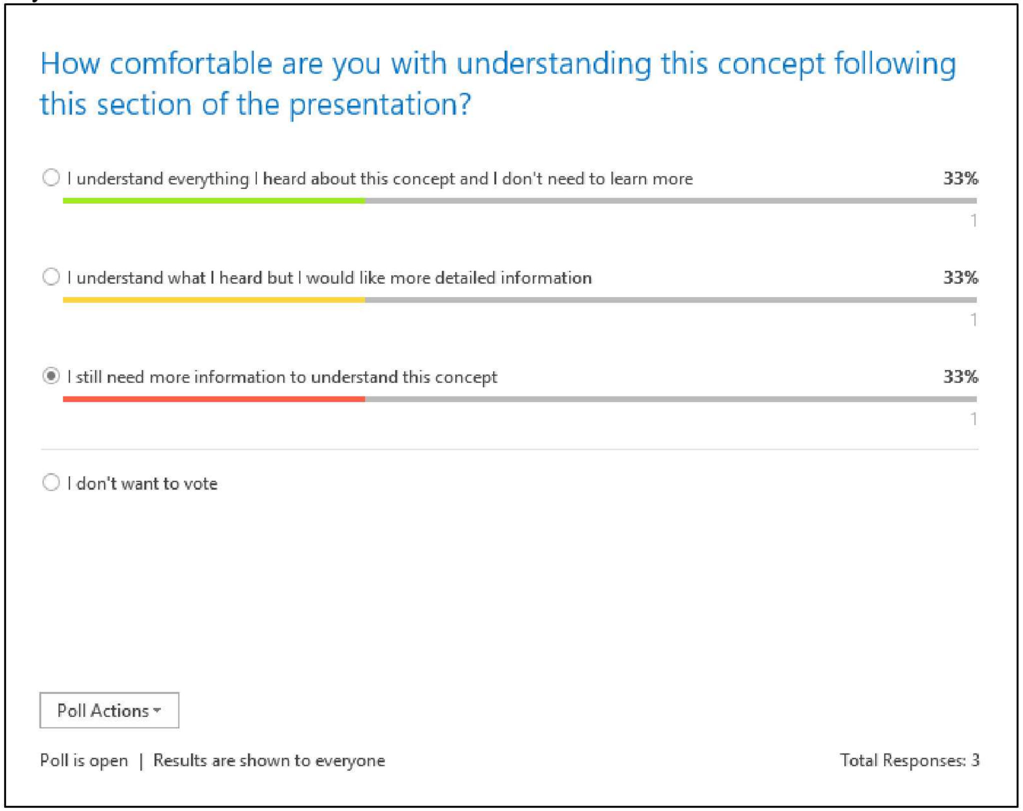
- UQ aggregation and roll-up
  - How have sources of uncertainty been combined and transferred across different levels of the system (i.e., validation hierarchy) and to the application domain?





Real-time Feedback

- Concepts:
  - Uncertainty quantification and sensitivity analysis credibility activities
  - Understanding how upstream credibility activities impact uncertainty quantification and uncertainty analysis results

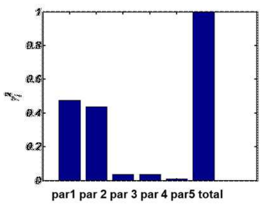


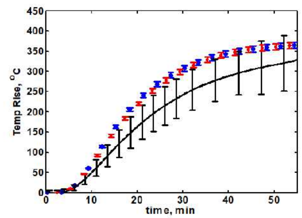
Credibility Evidence Supporting CompSim Predictions

Delivering Credibility Evidence with Final Results

- **Credibility evidence** can be delivered alongside predictions or in supporting documentation
- The **V&V/UQ/Credibility Assessment Communication Template Set** is designed for use in assembling and communicating the evidence-driven credibility aspects supporting CompSim predictions at **various levels of detail**.
  - Includes documentation of application context, credibility evidence details and/or summary, limitations and risks, key gaps, and potential path forward.

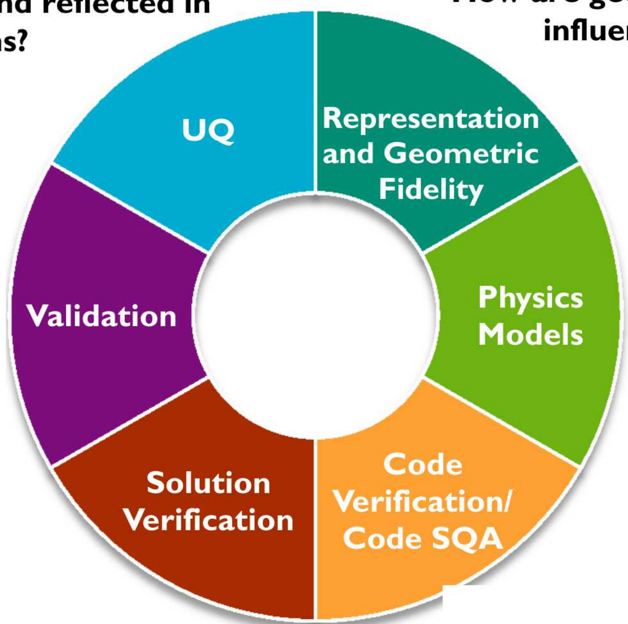
CompSim Credibility Evidence Summary

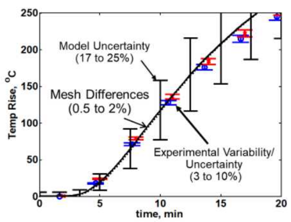


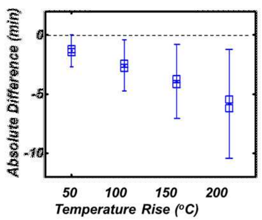


How are uncertainties assessed and reflected in simulation predictions?

How are geometric feature simplifications influencing simulation results?





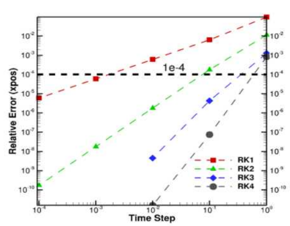


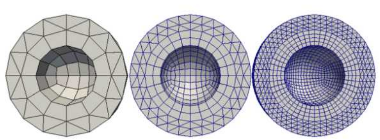
What is the discrepancy between simulation and experiments?

PIRT

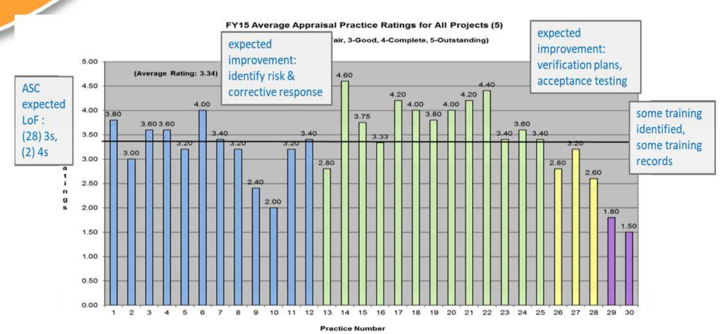
Phenomena	Importance	Adequacy for Intended Use			
		Math Model	Code	Validation	Model Parameter
Phenomena 1	H	H	M	L	L
Phenomena 2	M	H	M	L	L
Phenomena 3	L	H	M	L	L

Are important physics models adequate? Key gaps mitigated?





How do numerical solution or human errors affect simulation results?



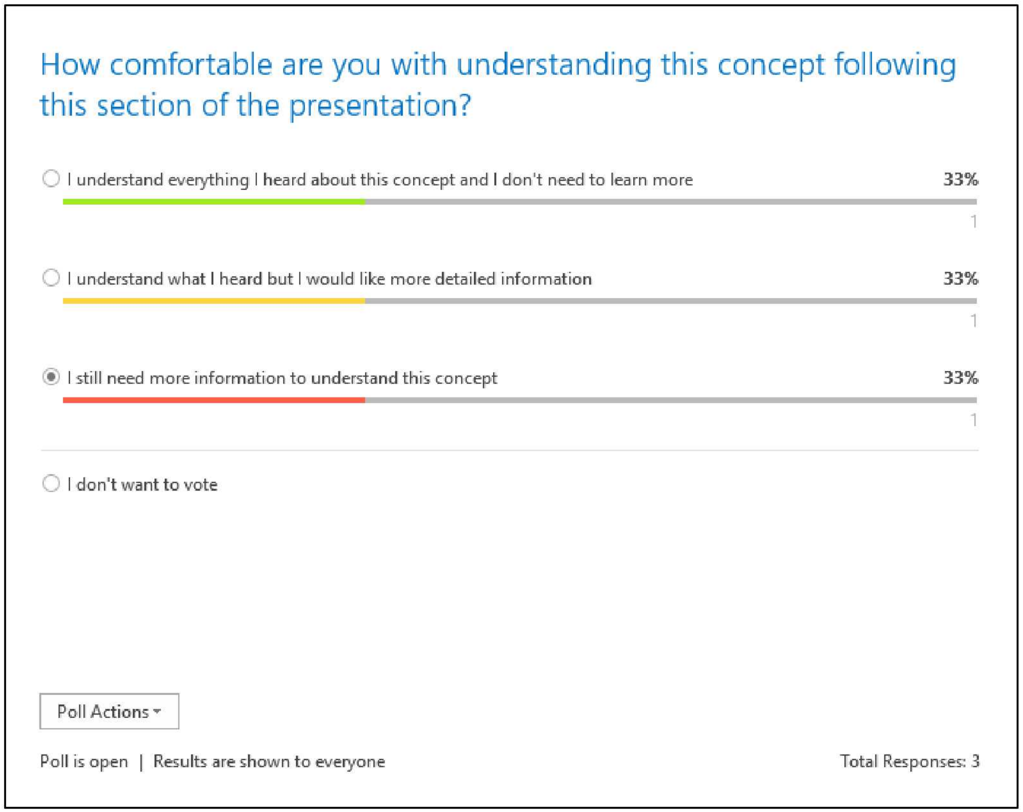
What is the evidence for code credibility?

The Credibility Process Provides Planning, Actions, & Results

- Credibility processes seek to support planning for, developing, and presenting **credibility evidence** to support CompSim predictions.
- Methods and tools continue to be developed – this process is not perfected and there is still work to be done.
- We continue to work towards better integration of V&V/UQ/Credibility into the CompSim workflow.

Real-time Feedback

- Concepts:
  - Compiling credibility evidence to support CompSim predictions



V&V/UQ/Credibility Resources

V&V/UQ Resources

- ASC and 1544 maintain the V&V/UQ portal: https://vvuq.sandia.gov/

- The ASC V&V/UQ Portal compliments the expertise of V&V subject matter experts by providing helpful core reference materials and tools for the SNL community. Key features include:
  - **Guidance Documents**, which are short 4-page summaries of processes and tools that give readers high-level introductions into "What concept is" or "How a tool can be used."
  - A **repository** that houses V&V/UQ related documents, reports, presentations, and other materials and serves as a reference archive that stores pertinent V&V information for future access.
  - A list of **contacts** who are available to answer any V&V/UQ question that may arise.

V&V/UQ Resources

- The ASC V&V/UQ Portal also includes links to **trainings and seminars**
  - ESP700 – "An Introduction to Verification, Validation and Uncertainty Quantification" – is available for streaming via the portal: https://vvuq.sandia.gov/esp700_previous#fy19
    - Session 1: Overview of V&V/UQ Concepts
    - Session 2: Code and Solution Verification
    - Session 3: Sensitivity Analysis, Uncertainty Quantification and DAKOTA Intro
    - Session 4: Validation of Computational Models and Course Wrap-up
      - Session 4 includes a full analysis of the projectile problem as an example
- The Statistics Department (6673) teaches classes throughout the year – see TEDS for offerings
- DAKOTA documentation and information is available on the DAKOTA website: https://dakota.sandia.gov/

Conclusions and Path Forward

Goal & Structure of this Training

- Goal
  - The purpose of this short course is to **introduce** V&V/UQ/Credibility process concepts, methods, and tools that have been developed for CompSim at Sandia.
- At this point we have covered:
  - Introduction to the CompSim Credibility Process at Sandia
    - Motivation and Historical Perspective
  - Integrating V&V/UQ/Credibility into CompSim
    - Mapping Requirements to a V&V/UQ/Credibility Strategy
    - Developing a V&V/UQ/Credibility Plan
    - PCMM Process Overview
  - PCMM Credibility Elements Overview
    - Geometric/Representation Fidelity
    - Physics Model Fidelity (PIRT)
    - Code Verification
    - Solution Verification
    - Validation
    - Uncertainty Quantification
  - Credibility Evidence Supporting CompSim Predictions
  - V&V/UQ/Credibility Resources
    - VVUQ Portal
    - Trainings
    - Guidance Documents

Wrap-up Discussion & Identification of Topics of Interest

- Are there any additional discussion points that we should cover?
- Please feel free to follow up with me with any additional questions