Credibility Framework – an End-to-End Credibility Workflow Platform

George E. Orient, Robert Clay (SNL)
Didier Verstraete (NexGen Analytics)

ASME V&V 2020 Symposium
May 20-22, 2020
Credibility Overview

Credibility Framework (CF) Requirements

CF Software Process

PIRT (Phenomena Identification and Ranking Table) Tool

PCMM (Predictive Capability Maturity Model) Tool

Summary, Current Work and Plans
Modeling and Simulation Credibility Process at Sandia

- The process of assembling and documenting *evidence* to ascertain and communicate the *believability* of *predictions* that are produced from computational simulations
- Quality process for CompSim (Computational Simulation)

**Application Context**
- Application Requirements
- Negotiate Role of CompSim in Decision Making
- Derived CompSim Requirements
- QoIs (Quantities of Interest)
- Test-CompSim Integration

**Planning and Execution**
- Model development and V&V
- Documentation
- Analysis governance
- Workforce qualification

**Assess & Communicate**
- Customer engagement
- Peer reviews
- Prediction issues
- Gaps and path forward

**Deliver Predictions**
- Plausible margin bounds
- Credibility evidence

**UQ**
**Validation**
**Solution Verification**
**Code Verification/Code SQA**
**Physics Models**
**Representation and Geometric Fidelity**

ND mission space: non-monotonic, discontinuous system responses - design and margin assessments under uncertainty REQUIRE agile execution of large model ensembles
Qualitative evidence
- SME judgment, tacit organizational knowledge, past history
- Expected predictiveness of the model for the intended use
- PIRT (Phenomena Identification and Ranking Table) - Defines key physical phenomena ranks their importance, identifies capability gaps
- Analysis governance, peer reviews

Quantitative “flavored” evidence
- PCMM (Predictive Capability Maturity Model) - SME elicitation process designed to characterize and communicate the completeness and rigor of the CompSim process.
- Quantitative elements such as UQ and Validation but aggregation is difficult

Validation at a handful conditions – mission space is large, response is nonlinear/discontinuous, test data are sparse

Need to combine qualitative and quantitative evidence to support decision making in large untested mission space
Aspirational effort to answer: Why should the customer believe predictions?

What is the risk of making decisions based on CompSim?
Credibility Overview

Credibility Framework (CF) Requirements

CF Software Process

PIRT (Phenomena Identification and Ranking Table) Tool

PCMM (Predictive Capability Maturity Model) Tool

Summary, Current Work and Plans
Tailor credibility process to match consequence of the CompSim predictions
  ◦ Trade studies in design support
    ◦ Quick turn-around, V&V trained analyst, input data starved, comparative
  ◦ CompSim based qualification
    ◦ Significant effort, dedicated V&V budget, up-front constitutive and subsystem tests, predictive
    ◦ Configurable by non-programmers through simple spreadsheets

Be flexible to adapt to organizational differences (PCMM, TRL, etc.)
  ◦ Credibility process elements and subelements vary
  ◦ If the organization/program requires then support gap analysis through assessment
    ◦ Acceptability of assessment while acknowledging metrics are not precise

Record different states throughout the lifecycle of the program

Support queries to identify important capability gaps

Integration with diverse data sources (SPDM, PLM, etc.) used for storing evidence

Auto-generating human readable credibility report distilled from vast data repositories

Open source effort aims to serve and be developed by diverse technical community
Credibility Overview

Credibility Framework (CF) Requirements

CF Software Process

PIRT (Phenomena Identification and Ranking Table) Tool

PCMM (Predictive Capability Maturity Model) Tool

Summary, Current Work and Plans
N-tiered:

- **Launcher:** used to start plugin and load configuration
- **Presentation:** contains plugin GUI
- **Business Logic:** contains business logic, coordinates plugin behavior, performs calculations, makes logical decisions and processes commands
- **Data:** queries persistent data from database or files
- **Model:** used to transmit data (in memory) to other layers
- **Tools:** contains functions and methods used across whole application

**Benefits:**
- Easy to manage
- Scalable,
- Flexible,
- Reusable
CF Software Process Elements

Main software requirements for CF plugin:

• Persistence through an open source database
• Configurable by non-programmers through familiar Excel spreadsheets

Implementation and software process details

◦ Eclipse plugin (Sandia Analysis Workbench or generic)
◦ Configuration files:
  ◦ PIRT and PCMM data schema are Excel files transformed to YAML
  ◦ Database:
    ◦ Use of Java Persistence API; EclipseLink for object-relational mapping (ORM) implementation
    ◦ HSQLDB used to locally store data into workspace (open source and developed in Java)
◦ Credibility file (.cf) file format:
  ◦ Single .cf file to store CF process data within the workspace
  ◦ A zip file of both database and configuration files (as done by e.g. Word):
    ◦ Easy to manage in SDM (Simulation Data Management)
◦ GitLab-based continuous integration
  ◦ Testing
  ◦ Unit tests: contained in separate plugin to test CF plugin features w/o including tests in installation package
  ◦ Integration tests performed with Maven Tycho
  ◦ SWTBot used to test the GUI (work in progress)
Shown docked in Sandia Analysis Workbench; also works with plain Eclipse
Credibility Overview

Credibility Framework (CF) Requirements

CF Software Process

PIRT (Phenomena Identification and Ranking Table) Tool

PCMM (Predictive Capability Maturity Model) Tool

Summary, Current Work and Plans
A Phenomena Identification and Ranking Table, or PIRT, provides a structured approach to identify and prioritize the important physical phenomena in an engineering application.

- Define **key physical phenomena** and rank their importance
- 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

A PIRT is developed through expert judgment for a particular intended use.
- The intended use is specific to the application driver, scenario, and analysis objective

Each QoI (Quantity of Interest) has its own PIRT
Application focused capability gap analysis; tracking history over project life cycle.
Credibility Overview

Credibility Framework (CF) Requirements

CF Software Process

PIRT (Phenomena Identification and Ranking Table) Tool

PCMM (Predictive Capability Maturity Model) Tool

Summary, Current Work and Plans
The Predictive Capability Maturity Model (PCMM) is a multi-dimensional qualitative metric to facilitate discussion and communication of credibility evidence

- Primary purposes:
  - Determine readiness of modeling capabilities and simulation products for use in various applications and decisions (e.g., design, environment specification, qualification)
  - 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 levels – a relative measure of the state of the evidence and level of effort around each element
  - Element criteria – major features of the evidence to consider for each element
  - Roles – who provided evidence and/or assessments? Customer, code developer, analyst, experimentalist, etc.
PCMM Elements

Code Verification

Analysis code reproduces closed-form results

Physics and Material Model Fidelity

Are “closure models” (constitutive etc.) credible?
E. g. MLEP (Multi-Linear Elastic-Plastic) WHY? Model form error?

Representation and Geometric Fidelity

Is the geometric abstraction acceptable?

Solution Verification

Code solves the equations for the intended use correctly?

Challenge: Often unsettling when modeling highly nonlinear, chaotic mechanical systems

Uncertainty Quantification

What is the effect of input uncertainties on QoIs?
- Uncertainty inventory and characterization of input uncertainties
- Formal UQ; propagate characterized uncertainties through the model
- Experimental uncertainty

Validation

Validation hierarchy
How well do model predictions match experimental data?
CF PCMM Configuration by Non-Programmers

Excel spreadsheets familiar to V&V practitioners

<table>
<thead>
<tr>
<th>Element</th>
<th>CVER</th>
<th>Code Verification</th>
</tr>
</thead>
<tbody>
<tr>
<td>PMMF</td>
<td>Physics and Material Model Fidelity</td>
<td></td>
</tr>
<tr>
<td>RGF</td>
<td>Representation and Geometric Fidelity</td>
<td></td>
</tr>
<tr>
<td>SVVER</td>
<td>Solution Verification</td>
<td></td>
</tr>
<tr>
<td>VAL</td>
<td>Validation</td>
<td></td>
</tr>
<tr>
<td>UQ</td>
<td>Uncertainty Quantification (UQ)</td>
<td></td>
</tr>
</tbody>
</table>

Solution Verification (SVER) & Return to Elements

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>Have an SQE process in place, discuss bugs/errors</td>
</tr>
<tr>
<td>Medium</td>
<td>Memo documenting/referencing the SQE process</td>
</tr>
<tr>
<td>High</td>
<td>Coordinate with code team on known deficiencies and status</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Level</th>
<th>Activities</th>
<th>Roles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>Evidence</td>
<td>Customer</td>
</tr>
<tr>
<td></td>
<td></td>
<td>System Engineer</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Analyst</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Experimentalist</td>
</tr>
<tr>
<td>Medium</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Solution Verification (SVER) & Return to Elements

<table>
<thead>
<tr>
<th>Level</th>
<th>Activities</th>
<th>Roles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>Evidence</td>
<td>Customer</td>
</tr>
<tr>
<td></td>
<td>Assess</td>
<td>System Engineer</td>
</tr>
<tr>
<td></td>
<td>Aggregate</td>
<td>Analyst</td>
</tr>
<tr>
<td></td>
<td>Stamp</td>
<td>Experimentalist</td>
</tr>
</tbody>
</table>

Levels

- Low Rigor
- High Rigor

Agile adaptivity to organizational requirements
Progress and role of the actor are recorded

Heuristic progress tracking

Role tracking

Tagging supports life cycle tracking and queries
CF PCMM Tool – Adding Evidence

Recommended folder structure contains artifacts employed as evidence generated
Evidence is opened with associated editor
### CF PCMM Tool – Assess (Optional)

#### CompSim Credibility Process

**Assess, PCMM > Solution Verification > Assess**

<table>
<thead>
<tr>
<th>Element/Subelement</th>
<th>Level Achieved</th>
<th>Evidence Links</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Physics Models</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PMMF1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PMMF2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PMMF3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PMMF4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Geometry Fidelity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RGF1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RGF2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RGF3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Solution Verification</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SVER1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SVER2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SVER3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SVER4</td>
<td>Verify simulation post-processor inputs decks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SVER5</td>
<td>Technical review of solution verification</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Validation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VAL1</td>
<td>Define a validation hierarchy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VAL2</td>
<td>Apply a validation hierarchy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VAL3</td>
<td>Quantity physical accuracy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VAL4</td>
<td>Validation domain vs. application domain</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VAL5</td>
<td>Technical review of validation</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>UQ</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UQ1</td>
<td>Aleatory and epistemic uncertainties identified and reported</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UQ2</td>
<td>Perform sensitivity analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UQ3</td>
<td>Quantify impact of uncertainties from UQ1 on quantitative outcomes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UQ4</td>
<td>UQ aggregation and roll-up</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Assess PCMM Subelement

- **Code:** SVER4
- **Subelement:** Verify simulation post-processor inputs decks
- **Level achieved:** Level 2
- **Comments:** Code developer team was engaged, and they provided a memo entered as evidence.

**Role:** Analyst

---

Role is associated with assessment
CF PCMM Tool – Aggregate (If Assessment Done)

Average assessment of multiple respondents; consensus but retaining diversity
CF PCMM Tool – Quality Stamp (If Assessment Done)

Simple visual representation of CompSim credibility evolution
Credibility Overview

Credibility Framework (CF) Requirements

CF Software Process

PIRT (Phenomena Identification and Ranking Table) Tool

PCMM (Predictive Capability Maturity Model) Tool

Summary, Current Work and Plans
Credibility Framework - Summary

On-going work
- UI/UX (User Interface/User Experience) testing started
- Early adoption on high consequence programs
- Engaging other organizations (KCNSC, different SNL ND programs) to test configurability of CF to match its behavior with their credibility process
- CF open source submission process started. Look for “Credibility Framework” on gitlab.com in October.

Plans (FY21 and beyond)
- Queries (PIRT: “What phenomena had ‘red’ gaps at the preliminary design review?”)
- Managing program requirements and evidence of meeting them
  - Interfacing with existing requirement management systems (DOORS, etc.)
- Experimental credibility
- Peer review framework
- Credibility risk management
- Credibility constructs at different consequence levels (design study, system test design, CompSim based qualification)
- Automated credibility report generation through
- Evidence theory (belief-plausibility) and UQ based verification of program requirements