FY05 Progress – From V&V to CUD

Tim Trucano
Optimization and Uncertainty Estimation Department, 9211
Sandia National Laboratories
Albuquerque, NM 87185

***Phone: 844-8812, FAX: 844-0918**
***Email: tgtruca@sandia.gov**

Sandia National Laboratories, 2005-3651P

Summary of accomplishments

- V&V
  - Program
    - Written program positions
    - V&V methods PI
      - ❑ Two technical papers
    - ASC, HQ support
  - HEDP Milestone
    - Broad support
  - Advocacy
    - Internal/external outreach
- CUD
  - Papers
  - Hook to Intelligence Community for Yonas
  - New slant on QMU
  - New connections to social models

I work with many talented people!

Advocacy of M&S

- **My real Mission:**

Advocate high-quality, high-impact computational science intelligently and constantly.

- **In other words, as Bill has put it: That M&S is VALUED by our Lab's national security culture, not just tolerated.**

- **V&V is one important way to do this, but not the only way.**

ASC V&V Program – Support Pilch

- It is important that I support Marty Pilch
- FY05 examples (written but not published):
  - "Roles and Responsibilities" (w/ Pilch) – used in ASC business model planning, ASC V&V strategic plan, local bridge building
  - "Comments on 'Responsiveness'" – analysis of V&V responsiveness issues (for Pilch)
  - "Comments on Progress Metrics: Tentative Ideas" – thinking paper supporting preparation for JASON "V&V" study
  - "Measures of Modeling and Simulation as a Sandia DP Way-of-Life" (w/ Pilch and Lott) – response to Sandia ASC strategic planning thrust about lack of understanding of large-scale M&S at Sandia (but what to do with it?)
  - Currently working on whitepaper on code verification metrics for JASON study (Pat Knupp also engaged on this)
- We are definitely tired of hearing the "too long … too expensive … too hard" stuff

JASON – JulyAugSeptOctNov

ASC V&V Program – Thinking

- PI of V&V methods ($3.76M in FY05) – cheerleader, coordinator, contributor:
  - Traditionally validation methods and UQ
  - Focused on supporting Lifetime Extension Programs (LEPs) work through tech development and prototyping
  - FY05 introduced verification into the mix, three projects, ALL in 9200 ($630K).
    - Temporal convergence focus (Ropp and Shadid)
    - Mathematics can really help (Bochev and Christon)
    - We need better test problems (Knupp & friends)
    - (Hope there is more of this in FY06*)
  - Two personal contributions:
    - "Calibration, Validation, and Sensitivity Analysis: What's What" (w/ Swiler, Igusa, Oberkampf, Pilch) – accepted for Reliability Engineering and System Safety
    - "Software Engineering Intersections with Verification and Validation (V&V) of High Performance Computational Science Software: Some Observations" (w/ Post, Oberkampf and Pilch) – R&A for electronic posting
  - Still working on verification ideas and Quantitative Margins and Uncertainty (QMU)

*Caveat: nobody knows impact of 15-20% potential budget cut.

HEDP (High Energy Density Physics Level II Validation Milestone – Q4 FY05

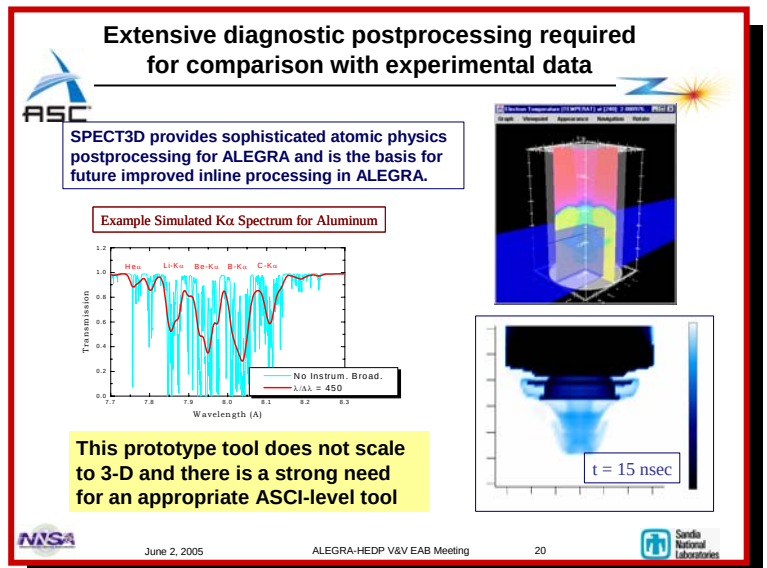
- Tom Mehlhorn owns the operational delivery of the milestone
- Team effort between 1600 (HEDP Theory) and 9200 (ALEGRA)
- I think, advise, review and criticize, and actively contribute to the milestone team on specific milestone elements:
  - "Comments on Validation" for internal purposes
  - Validation plan in progress
  - Consulting on verification in progress
- I also advise and provide progress feedback to Pilch
- Expectation* is that HEDP will remain in V&V funding stream in 06

*Caveat: nobody knows impact of 15-20% potential budget cut. HEDP was one of five year thrust areas specified by ASC strategic planning in early 2005.

Milestone Specification Summary

VV-II-5.1 Deliverable	YES	NO
1. Present a validation plan that guides this milestone work, reviewed and approved by Sandia Z-program management.	60%	
2. Assess code suitability for validation activities	90%	
2a. Provide evidence of code documentation (theory, algorithm and software documents, such as user manuals)	90%	
2b. Provide evidence that codes are managed to institutional SQA requirements	~100%	
2c. Provide evidence of code verification	~50%	
2d. Check code issue log prior to validation simulations	~100%	
3. Characterization and validation of EOS, Opacity and Resistivity models	~60%	
3a. Evidence of community standard material data	~50%	
3b. Analysis of simulation sensitivity to material data	~90%	
3c. Comparisons with material data with material experimental data	~70%	
3d. Evidence of material data validation from at least one non-Z-machine experiments	~25%	
4. Non-Z-Machine Validation Data Study	0%	
4a. Present the test matrix		
4b. Evaluate solution sensitivity to discretization parameters		
4c. Quantify measurement/prediction comparisons		
4d. Assess credibility for the intended application		
5. Z-Machine Magnetically Launched Flyer Study	~80%	
5a. Present the test matrix	~90%	
5b. Evaluate solution sensitivity to discretization parameters	~90%	
5c. Quantify measurement/prediction comparisons	~80%	
5d. Assess credibility for the intended application	~80%	
6. Z-Machine Dynamic Hohlraum Study	~10%	
6a. Present the test matrix	~50%	
6b. Exploration of potential methods to understand sensitivity of these calculations to ALEGRA-HEDP numerical discretizations (solution verification)	~0%	
6c. Initial evaluation of possible measurement/prediction comparisons using specified validation metrics	~0%	
6d. Specification of relevant/important physics R&D issues	~0%	
7. Z-Machine Bare Pinch Study	0%	
7a. Present the test matrix		
7b. Exploration of potential methods to understand sensitivity of these calculations to ALEGRA-HEDP numerical discretizations (solution verification)		
7c. Initial evaluation of possible measurement/prediction comparisons using specified validation metrics		
7d. Specification of relevant/important physics R&D issues		
8. Archive available documentation	~50%	

- **This validation milestone is different**
  - **V&V in R&D focused project, i.e. more science, less LEP**
  - **Encompass frequent technical innovation**
  - **Looks to the future**
  - **But still a bottom line somewhere near 2010 – the new machine key decisions**



National Laboratories

Outreach and Advocacy

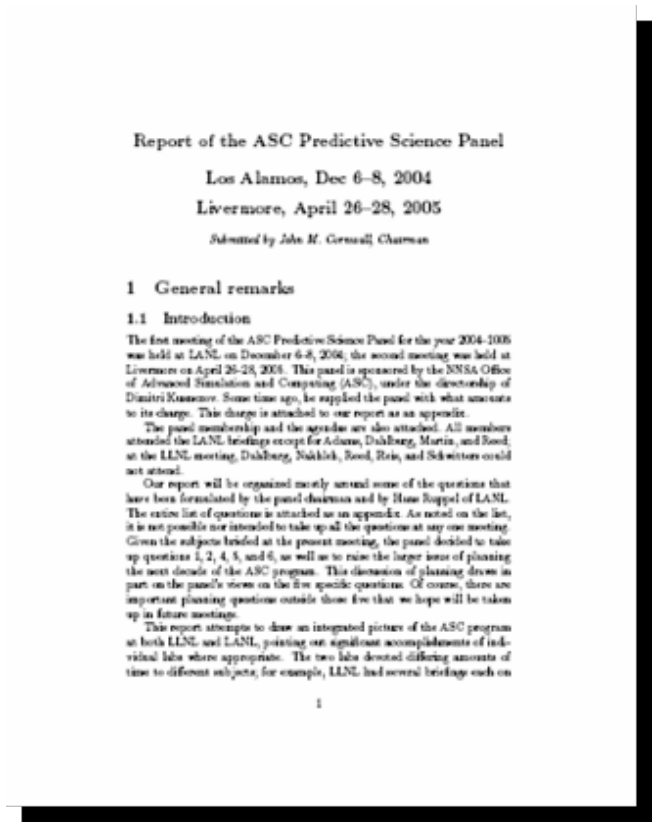
- **<u>It is also important that I support the Tri-Lab V&V Program and more generally the ASC program</u>**
- **ASC V&V Ten-Year Strategy – I worked directly with HQ:**
  - **"Future Strategy for V&V – Talking Points"**
  - **"The Verification and Validation Program Strategy" (3/16/05 version)**
- **Participation on ASC-related advisory and milestone review panels.**
  - **Predictive Science Panel (ERA Award)**
  - **Three LANL-related review panels (one FY04 milestone delayed six months; one FY05 milestone; one LDRD review)**
- **Sandia outreach and consulting:**
  - **QASPR (working with Hutchinson as closely as he needs)**
  - **On-call at Sandia for anybody**
- **LANL (and LLNL if they wish) outreach and consulting:**
  - **LANL designer(s) talking to me (with encouragement of Peery and Hommert)**

Example of OUTREACH: Experimental data should have error bars!

Predictive Science Panel – Service to ASC and NNSA

- My role: Member of the Panel (requested by HQ and both LANL/LLNL)
  - I bring V&V, breadth of knowledge, sanity
  - I look out for Sandia
- Let me emphasize: it is called the "ASC Predictive Science Panel" – they visit only LANL and LLNL.
- HQ sees links between:
  - "Predictive Science"
  - QMU
  - V&V
- Read the ASC strategic vision!
- Camp and Hale are engaged

FY05 Final Report



Report of the ASC Predictive Science Panel

Los Alamos, Dec 6–8, 2004

Livermore, April 26–28, 2005

Submitted by John R. Cornwall, Chairman

1 General remarks

1.1 Introduction

The first meeting of the ASC Predictive Science Panel for the year 2004-2005 was held at LANL on December 6-8, 2004; the second meeting was held at Livermore on April 26-28, 2005. This panel is sponsored by the NNSA Office of Advanced Simulation and Computing (ASC), under the directorship of Dimitri Kusnezov. Some time ago, he supplied the panel with what amounts to its charge. This charge is attached to our report as an appendix.

The panel membership and the agendas are also attached. All members attended the LANL briefings except for Adams, Dahlburg, Martin, and Reed; at the LLNL meeting, Dahlburg, Nakhleh, Reed, Reis, and Schwertfeger could not attend.

Our report will be organized mostly around some of the questions that have been formulated by the panel chairman and by Hans Ruppel of LANL. The entire list of questions is attached as an appendix. As noted on the list, it is not possible nor intended to take up all the questions at any one meeting. Given the subjects briefed at the present meeting, the panel decided to take up questions 1, 2, 4, 5, and 6, as well as to raise the larger issue of planning the next decade of the ASC program. This discussion of planning draws in part on the panel's views on the five specific questions. Of course, there are important planning questions outside those five that we hope will be taken up in future meetings.

This report attempts to draw an integrated picture of the ASC program at both LLNL and LANL, pointing out significant accomplishments of individual labs where appropriate. The two labs devoted differing amounts of time to different subjects; for example, LLNL had several briefings each on

1

Sandia National Laboratories

JASON 2005 Summer Study on Exactly What? June 2005

- Initially, topic was V&V and associated progress measures
- Evolved to “progress measures” for the ASC program
- Now defined as “Predictive Science” progress measures and “brainstorming”

DANGER!

- Sandia initially not invited to present
  - “Ideas from LANL and LLNL will cover Sandia.” (Kusnezov)
- Now Sandia is on agenda, but is the damage going to multiply?

Advanced Concepts Group: CUD – Cognition and Uncertainty in complex Decisions

CUD – How did this happen?

Interaction map

6/04
UQ for DHS Whitepaper

8/04
Laura McNamara Stewardship Study

10/04
Laura McNamara In ACG

Changed My Life!

8/04
w/ 9216 CSRF Proposal for Agent V&V

11/04
ACG Funds me To Work with Laura And on UQ/V&V For social models

12/04
Having contact With Jessica Turnley On V&V

Continued interest in UQ for DHS but … where to go and what to do?

12/04
Laura/Tim on I2K2W Team – Intelligence Analysis Emphasis

6/05
Maintaining Contact w/ 9216 On agent V&V

6/05
Laura Working on L-M LANL Bid Team

???

> FY05: Turnley, 9216 (Jorgensen, Sprigg, Backus, Boslough, …)

CUD – Why do I (Laura, Jessica) care?

- First, CUD is isomorphic to decision making under uncertainty problems that I've become very interested in
  - In particular, QMU sees itself in this world
- Second, this is a way of looking at M&S and decision making in social science – which I'm interested in
- Third, the UQ challenges in non-physical science are extraordinarily hard; there is the opportunity to learn a lot
- Fourth, we have discovered that a little thinking goes a long way in these circles
- Fifth, this is a way of bringing V&V back to problems of interest to 9216.
- Sixth, I want to have some <u>impact</u> in ACG.

CUD – What have I (w/ Laura and Jessica) done?

- First, I2K2W → CUD as the official acronym
- Second, Laura and I introduced a spectrum of canonical decision areas that tracks the transition from intuitive to analytical judgment CUD is moving on (see below)
- Third, I (primarily) introduced the notion of an "Intelligence Skunk Works" as a focus for the CUD team work on the U.S. Intelligence Community
  - One application of the concept of organizational sensemaking (Laura provided the traction on this)
- I'm providing general thinking to ACG on UQ and V&V issues
  - Example 1: CIA <u>isn't ready</u> for sophisticated UQ technology even if it existed
  - Example 2: Retrospective understanding of information is critical; this <u>is</u> a good place to insert technology
  - Example 3: When three-letter agencies appear I come when Gerry requests

CUD Spectrum: Analytic UQ to the LENS Model

Intuitive Emphasis ⟷ **Analytic Emphasis**

Intelligence Analysis	Economic Forecasting	Climate Forecasting	Stockpile Stewardship
CUD Team FY05 Emphasis Boslough asymmetric info aggregation/UQ	9216 V&V insertion + George Backus decision problems	Tim – previous experience; interesting V&V extension	Laura McNamara project w/ Tim

Work on social model V&V insertion

CUD – products

- ACG Newsletter w/ Laura on SKUNK - published
- ACG Newsletter “From Pencils to Computers” – trying to publish
- Sandia Report: “Skunk Works as a Metaphor for Organizational Design in the Intelligence Community: A Thinking Paper” in R&A (w/ McNamara, Turnley, Whitley, Chew)
- Some steps to mapping ASC V&V concepts to social models
  - Commentary on Turnley’s Whitepaper on validation in social science
  - Presentation to agent-based modeling tech exchange
  - Writing paper with Turnley on social model V&V
- Realization that Judgment Theory is the proper language to express the assertion component in stockpile QMU (!!)

(Remember: for this stuff all Gerry cares about is simple straightforward short viewgraphs)

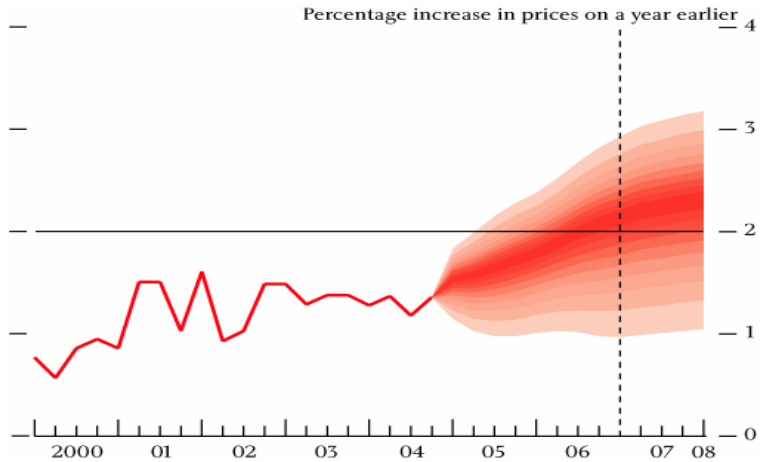
CUD – Where will this go?

- I would like to continue to get ACG funding and especially work with Jessica Turnley, who really knows what she is doing
  - May get ACG LDRD for a validation experiment study w/ Slepoy and Jessica
  - But I want more funding; and I don't want to move to ACG to get it
- When L-M owns LANL, Laura McNamara and I will take up the Stockpile Stewardship project again
- This year was a launching pad to economics-like decision processes
  - I would like to start talking to George Backus a lot more
  - Yonas has already asked for us to do an ACG brainstorm on economic forecasting challenges
- Uncertain information aggregation – what underlies Boslough's very smart idea about using markets as info aggregators for CUD?
- Can we get anywhere with Intelligence production and use? CUD issues are huge in the Intelligence Community!
- Can we get anywhere with DHS?
- What are opportunities for 9200?

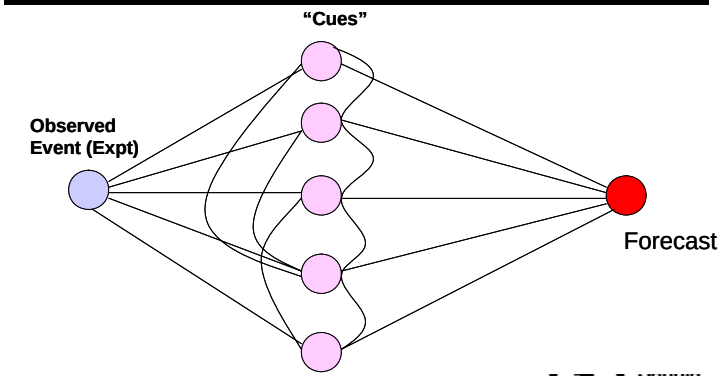
Social Models – it's not all cows calmly chewing

- Keep in mind:
  - How to do UQ and V&V? is less important than …
  - Does the community want UQ and V&V?
- Assuming an appropriate rationale for evaluation (a BIG if!) there are plenty of ideas available:
  - Evaluation criterion based on quantified "satisficing" (multiobjective OUU)
  - Application of Information-Gap theory for UQ
  - Use of elicitation methods (Booker, et al) for communication of uncertainty as well as quantification
- Even more interesting potential ideas:
  - Use of Qualitative Choice Theory (Backus)
  - Asymmetric information (including information aggregation and dynamics)
  - Quantified judgment theory (skill score development; cue uncertainty; cognitive correlations)

Ex: Economic forecasting



LENS Model



Miscellaneous

- **Cheerleading on PRIDE.**
  - **This project should be renewed. All the hoops ES told us to jump through were done.**
- **Cheerleading Calibration Under Uncertainty/Bayesian methods with Laura Swiler**
  - **The RESS paper contributes**
  - **An old product – review of Calibration Under Uncertainty literature - emanating from MICS funding in FY04 finally ready going to be published (Laura Swiler did 95% of the work)**
  - **This work should have a life of its own at this point: many interesting problems, impact on PRIDE, Pilch supporting, a natural intellectual framework to express model discrepancies sampled by V&V activities**