The 2014 Sandia V&V Challenge Workshop and the ASME VV&UQ Journal

Ken Hu

V&V, UQ, and Credibility Processes Dept.
Sandia National Laboratories

Presented at the 2015 ASME V&V Symposium
Thanks up front

- Challenge workshop participants
- V&V department and Dakota team at Sandia
- ASME committees
- Symposium organizers
- V&V community
Should we have a workshop?

- Of course! But think about:
  - Why
  - Who
  - What = workshop with challenge problem
  - How
  - Where = ASME V&V Symposium
  - When = 2011~2014+

- Start with why: be very clear on the goals
  - but you also need participants to come.
  - and they must be able to complete the problem
Be very clear on the goal

Pick one?

- Community engagement
- Education/training
- Methods demonstration
- Solve a real world problem
- Methods development
- Discuss an open problem

Our goal evolved over two years of development

- Wanted to provoke questions about role of V&V
- The resulting challenge problem lacks focus
- Can apply almost any V&V related question
Be very clear on the goal

Started in late 2011, same timeframe as this symposium
- Initial goal: highlight V&V topics of interest to Sandia
- How to differentiate from a conference?
- Build around a challenge problem
Be very clear on the goal

- Provide a venue for analysts to learn best practices
  - Provide valuable experience
  - Can be done without Sandia
  - Ex: short courses, literature, standards committees

Community engagement

Education/training

Methods demonstration

Solve a real world problem

Methods development

Discuss an open problem
Be very clear on the goal

- Community engagement
- Education/training
- Methods demonstration
- Solve a real world problem
- Methods development
- Discuss an open problem

Continue the tradition of challenge problems

- Check the state of the art in V&V methodology
- Focus on one V&V topic, ex: “aggregation”
Be very clear on the goal

Community engagement

Education/training

Methods demonstration

Solve a real world problem

Methods development

Discuss an open problem

Crowdsourse the solution to a real problem

- Ex: Netflix prize
- Not a good fit for this community
- Real problems too big, complex, proprietary
Be very clear on the goal

Targeted development of methods or approaches

- Ex: extrapolation out of validation domain
- Needs time, funding
Be very clear on the goal

Workshops are good for highlighting gaps

- Not many gaps on quantitative methods side
- Many open problems on qualitative process side
  - Ex: connection between V&V and credibility
Be very clear on the goals

Community engagement
Education/training
Methods demonstration
Solve a real world problem
Methods development
Discuss an open problem

We ended up with several goals, large scope

- Attract wider audience – something for everyone
- Less focused, more difficult for participants
Who: the V&V community

- Community is segmented in many ways
- At ASME, separate physics disciplines/ business sectors
  - V&V 10 vs. 20, 30, 40
- Subfields: V vs. V vs. UQ
  - Can’t be an expert in everything
- Different roles:

  Researchers  Practitioners  Customers
V&V community: Roles viewpoint

- R&D presence is strong
- Practitioners group is large, but not as visible

Researchers → V&V Theory
Practitioners → V&V Evidence
Customers → Requirements

V&V Theory
V&V Methods
V&V Standards
V&V community: “Roles viewpoint”

- Who are the customers?
- Where are they?
Who is the Workshop for

- Highlight what we can do
- Discuss the value & impact
- Imagine the customers
What? How?

- Pose a challenge problem, host a workshop
- Why and who → define the challenge problem

1. Aggregation of uncertainty
   → need many sources of uncertainty

2. Investigate the role of V&V
   → need V&V evidence to evaluate
   → need “real world” context

3. Accessible to many
   → multiple points of interest
   → wide range of possible approaches
   → limit barriers to entry
The Story of Mystery Liquid Company

Have many storage tanks, holding Mystery Liquid under pressure

During standard safety testing, one tank's measurements exceeded a safety specification

How should we respond?
Are the tanks at risk of failure?
No tanks have actually failed, ever.

Experimental and modeling efforts are begun
Supply a prediction – is it credible?

- How will evidence from experiments and simulations be integrated and used to support the final decision?

**What analyses?**

1. Characterize uncertainty from data
2. Treat epistemic vs. aleatoric uncertainty
3. Calibrate model parameters
4. Sensitivity analysis,
5. Parametric uncertainty quantification
6. Solution verification
7. Validation
8. Aggregation of uncertainty
9. Assess relevancy on a hierarchy of info
10. Assess credibility
The Challenge:

1) Develop and communicate a strategy to use experimental data and models

2) Predict failure probability at max load and account for uncertainty

3) Assess prediction credibility
ASME VV&UQ Journal

First issue will be dedicated to the challenge workshop

- Aniruddha Choudhary, I. Voyles, C. Roy, M. Patil (Virginia Tech), B. Oberkampf (Consultant)
- Zhimin Xi (University of Michigan – Dearborn), R. Yang (Ford)
- Lauren Beghini, P. Hough (Sandia National Labs)
- Tom Paez, P. Paez, T. Hasselman (Consultants, V&V10)
- Wei Chen, W. Li, S. Chen, Z. Jiang (Northwestern)
- Josh Mullins, S. Mahadevan (Vanderbilt)
- Michael Shields* (Johns Hopkins)
- Additional papers: intro, problem statement, truth model description, conclusion

* Participated, but no paper
Responses and reactions

- Groups applied different V&V strategies, methods
  - Reflects different priorities, time commitments
  - 6 from the R&D side, 1 group of practitioners
- \( P(fail) \) target is < 0.001
- Results for groups A-F
  - (A) 0.0075, (B) 0.0068, w/ high uncertainty
  - (C) Bounded by [0, 0.0034]
  - (D) 5e-16, with 99% confidence
  - (E) 0, with low simulation credibility
  - (F) N/A, data too poor to provide a prediction
Responses and reactions

- Very pleased with the responses
- Large audience, great discussions at the workshop
- There is no “right” answer
  - Six “valid” responses to the same challenge
  - Diversity in methods and strategies
  - Different ideas of what the V&V product looks like
  - Very different results
- Completed part 1 – methods demonstration
What’s Next

- How to evaluate / compare these results?
  - To each other & to the “truth” model
- Implications of very different results?
- How to choose a V&V strategy?
- How to assess credibility? Is it discipline dependent?
- What are the roles of V&V analyst and customer?
- How does V&V evidence influence decisions?
- We are now soliciting discussion papers for a later issue of the ASME VV&UQ Journal
Did we achieve our goals

- Community engagement
- Education/training
- Methods demonstration

- Solve a real world problem
- Methods development
- Discuss an open problem

Did we reach the right people

- Researchers
- Practitioners
- Customers

Halfway there
Thanks!

- I left out:
  - Sandia’s mission → unique V&V perspective and workshop motivation
  - Why Symposium was a good fit for the workshop
  - When (the timeline)
  - History of V&V/UQ challenge problems
  - Responses to the problem → Another talk, later today
- To learn more about the problem, workshop
  - https://share.sandia.gov/vvcw
  - Email us: vvcw@sandia.gov