Perspectives on Verification, Validation, and Uncertainty Quantification

William L. Oberkampf, PhD
Consultant
Albuquerque, New Mexico
wloconsulting@gmail.com

SIAM Conference on Computational Science and Engineering
Miami Hilton Hotel
Miami, Florida
March 2 - 6, 2009
Outline of the Presentation

• Uses of computational simulation
• Verification, validation and uncertainty quantification
• Where do we stand?
• Research and implementation issues
• Closing Remarks

Work in collaboration with Tim Trucano and Martin Pilch, Sandia Nat'l. Labs., and Scott Ferson and Jon Helton, consultants.
Typical Research Activity in Computational Science and Engineering
Uncertainty Quantification Included in Analyses for Decision Making
Verification and Validation Included in High-Consequence Decision Making
Verification Activities

• Definition used by U.S. DoD, AIAA, and ASME:
  Verification: The process of determining that a model implementation accurately represents the developer’s conceptual description of the model and the solution to the model.

• Two elements of verification are well recognized:

  • Code Verification: Verification activities directed toward:
    – Finding and removing mistakes in the source code
    – Finding and removing errors or weaknesses in the numerical algorithms
    – Improved software reliability using software quality assurance practices

  • Solution Verification: Verification activities directed toward:
    – Assuring the appropriateness of input and output data for the problem of interest
    – Estimating the numerical solution error, e.g. error due to finite element mesh resolution and time discretization
Validation Activities

• Definition used by U.S. DoD, AIAA, and ASME:
  
  **Validation**: The process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model.

• Validation is concerned with three activities:
  – Model accuracy assessment by comparison with a referent
  – Application of the model to the intended use, e.g., conditions where no referent data exist
  – Decision of model adequacy for the intended use

• Engineering and science communities require that the referent be **experimentally measured data**

• DoD allows **any** reasonable referent

• IEEE and ISO use different definitions of V&V, but they can be viewed as more general definitions
Uncertainty Quantification Activities

- Key sources of uncertainty:
  - Identification of environments and scenarios of the system
  - Input uncertainties in the system and in the surroundings
  - Model form uncertainty, i.e., uncertainty in $f(\cdot)$
Where Do We Stand?
Verification Activities

• Code verification:
  – Some commercial codes have extensive test suites composed of traditional analytical solutions
  – Weaknesses with code testing:
    • Traditional analytical solutions do not test complex coupling of terms
    • Order-or-accuracy testing is not done
  – Government, corporate, and university code testing is spotty, at best

• Software quality assurance (Hatton, 1997):
  “Scientific calculations should be treated with the same measure of disbelief researchers have for unconfirmed physical experiments.”

• Solution verification:
  – Error estimation usually relies on experience of the analyst, instead of quantitative error estimation
  – If model predictions agree with experimental data, there is little enthusiasm for investigating possible numerical errors
  – Sometimes it is fully recognized that numerical errors are as large as physics modeling errors, so model parameters are calibrated to adjust
Where Do We Stand: Validation Activities

• Common approach to validation is actually model calibration:
  – Parameters in the model, either scalars or probability distributions, are adjusted so that the model agrees with the experimental data
  – Usually reliable when the models are used for very similar systems and conditions where the models are calibrated
  – Weaknesses in the models, or coding errors, are rarely uncovered

• A relatively new approach to validation:
  – Emphasis is on assessment of model prediction inaccuracy, in the sense of a blind-prediction
  – Quantitative measures of disagreement (validation metrics) are assessed between model predictions and experimental measurements
  – More reliable when using the model to predict system responses:
    • Far from the conditions of the validation experiments
    • When the complete system can not be tested
Where Do We Stand: Uncertainty Quantification Activities

• Approach used in most high-consequence systems:
  – Characterize all uncertainties as either aleatory or epistemic:
    • Aleatory: inherent variation associated with the quantity, represented as a probability distribution
    • Epistemic: uncertainty due to lack of knowledge of the quantity, represented as an interval
  – Propagate input uncertainties through the model using Monte Carlo sampling techniques
  – Use alternate models to investigate model form uncertainty

• Bayesian approach:
  – Assume prior distributions for uncertain parameters in the model
  – Update the prior distributions for uncertain parameters using available experimental data an Bayes formula
  – Use Monte Carlo sampling, MCMC, or construct surrogate models to propagate uncertainties and update prior distributions
  – Compute new predictions using updated parameter distributions
Research and Implementation Issues: Verification Activities

• Develop manufactured solutions for a wide range of physics and engineering disciplines for order of accuracy testing

• Develop improved measures of code coverage in testing software; line coverage in regression testing is inadequate

• Develop less expensive and more robust methods for estimating spatial and temporal discretization error

• Develop numerical error estimators for nonlinear parabolic and hyperbolic PDEs

• Require improved code verification evidence from code developers

“I’ve already refined the mesh down to the microstructure of the metal!”
Research and Implementation Issues: Validation Activities

• Improve coordination and synergism between experimentalists and computationalists in designing and executing validation experiments

• Develop consortia to share validation test data among industry, commercial software companies, government, and universities

• Develop improved validation metrics to deal with:
  – Epistemic uncertainty in either the model or the experiment
  – Time series analysis

• Using the Bayesian updating approach, improve the separation of parameter updating and model error estimation

“Our results agree with the experimental data, why are you being difficult?”
Area Validation Metric

• The validation metric is defined to be the area between the CDF from the simulation and the empirical distribution function (EDF) from the experiment

\[ d(F, S_n) = \int_{-\infty}^{\infty} |F(x) - S_n(x)| \, dx \quad \text{(Minkowski L}_1\text{ metric)} \]
Research and Implementation Issues: Uncertainty Quantification Activities

• Improve the recognition and interpretation of aleatory and epistemic uncertainty

• Conduct further research and application of:
  – Probability bounds analysis (second order analysis)
  – Evidence theory (Dempster-Shafer theory)

• Extend Bayesian methods and polynomial chaos methods to incorporate interval-valued quantities

• Develop improved methods for estimating the change in model form uncertainty due to extrapolation:
  – Construct a non-Euclidian space for extrapolation
  – Map system response quantities to a probability space and then use the model prediction as an inverse transform to return to physical space

• Develop improved methods for sensitivity analysis when uncertainties are both aleatory and epistemic in nature
Effect of Characterizing Epistemic Uncertainties as Intervals versus Uniform Distributions

![Graph showing cumulative probability against system response quantity, comparing interval-valued responses to those representing epistemic uncertainties as uniform distributions.](image)
Risk-Informed Decision Making

- Nuclear power plant safety
- New aircraft design
- Scientific research
- Oil exploration
- Underground storage of nuclear waste
- Financial investing
- One week weather forecasting
- One year weather forecasting
- Human impact on global warming
Closing Remarks

• Verification and validation are processes that develop evidence of credibility in simulations

• Uncertainty quantification should forthrightly estimate:
  – Uncertainty associated with identified environments and scenarios
  – Uncertainty in simulation input quantities
  – Uncertainty in the model form applied at the conditions of interest

• What can be learned from failures in computational simulation?
  – Weaknesses in identifying failure modes
  – Under estimation of both aleatory and epistemic uncertainty
  – Inadequate quantification of model form uncertainty
  – Ability of decision makers to influence the analysis outcomes

V&V&UQ are concerned with truth in simulation, not marketing.

• We must recognize that engineering analysis, and how it is coupled to decision making, has fundamentally changed.
Some Prefer to Take the Position

“I don’t have the time, money, or people to do V&V&UQ.”
Suggested References


wloconsulting@gmail.com
Suggested References


Suggested References


