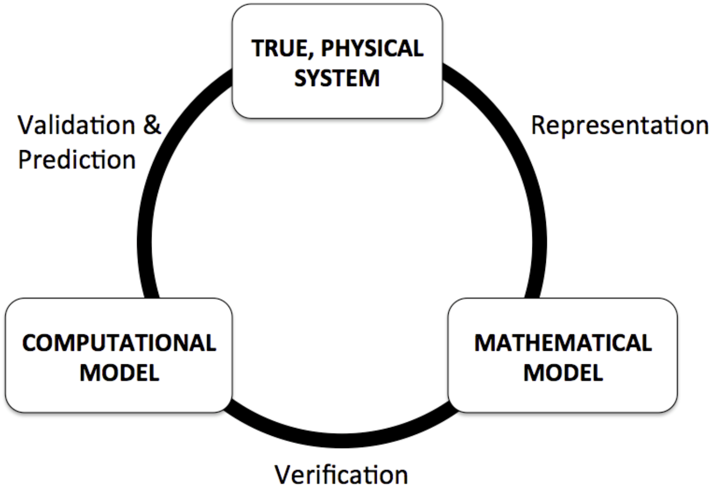
VVUQ: Principles and Best Practices

From NRC report "*Assessing the Reliability of Complex Models: Mathematical and Statistical Foundations of Verification, Validation, and Uncertainty Quantification*"

March 28, 2012

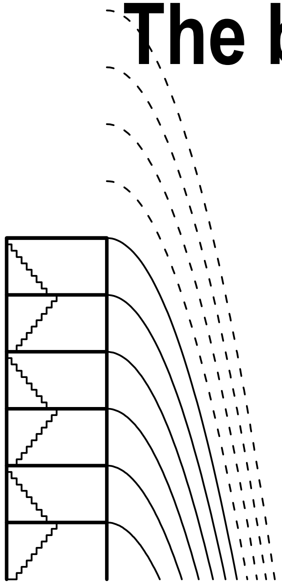
Marvin L. Adams (Co-Chair)
Texas A&M University

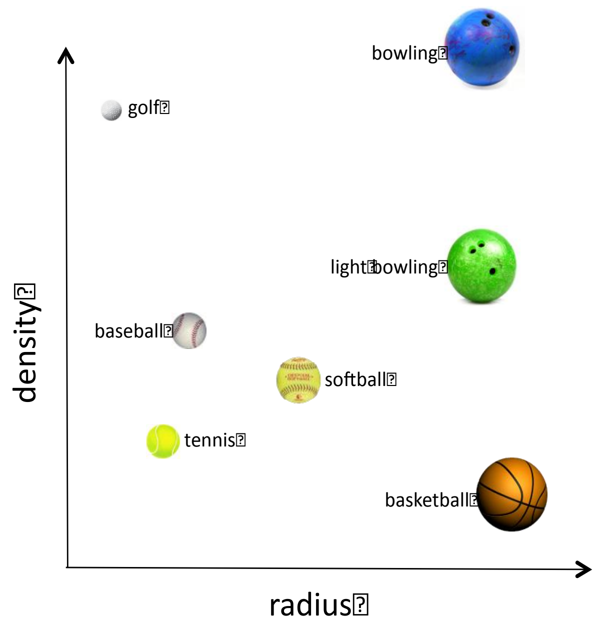
Set the stage: fundamental concepts and terms



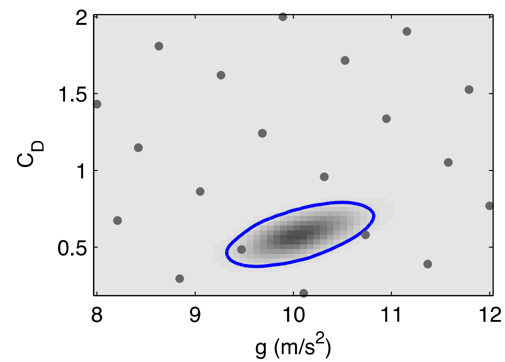
- **Math model:** math representation of reality (always approximate)
- **Comp model:** implements algorithms for approx. solution of math model
- **Verification**
  - Code: find and eliminate bugs
  - Solution: quantify numerical error
- **Calibration:** use data to infer values of uncertain parameters
- **Validation:** assess accuracy of model for its intended use
- **Prediction:** prediction of a QOI from a "new" problem
- **Uncertainty Quantification (UQ):** quantify range of values that a QOI may take in a given problem (includes propagation of input uncertainties)

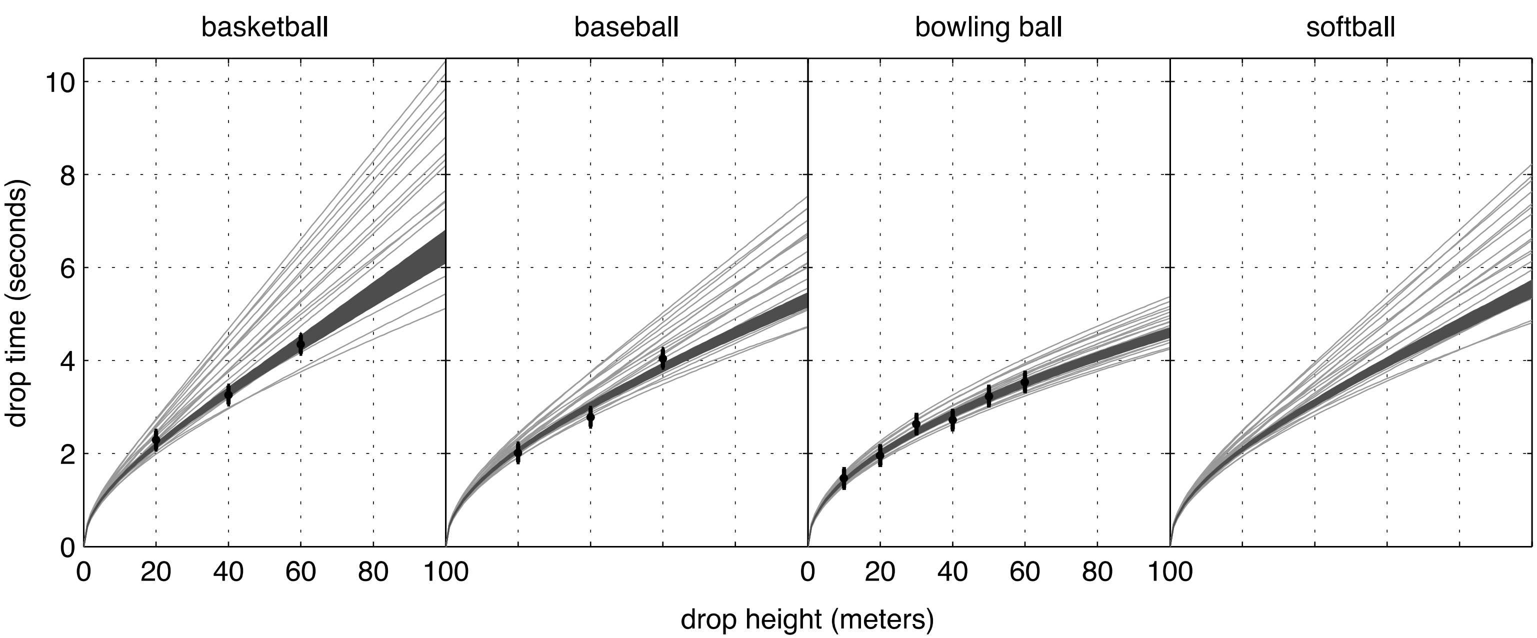
The ball-drop example illustrates much of this.





$$\frac{d^2h}{dt^2} = g - \frac{C_D}{2}\frac{3\rho_{\mathrm{air}}}{4R_{\mathrm{ball}}\rho_{\mathrm{ball}}}\left(\frac{dh}{dt}\right)^2$$





We often speak in terms of the “problem at hand.”

- We envision a setting in which specific QOIs must be predicted for a specific problem.
  - Example: electricity generated as a function of wind speed (the QOI) by a specific proposed wind-turbine design (the “problem”)
  - Example: drop time (the QOI) of a golf ball from 100m (the “problem”)
- The combination of specific QOIs and the specific problem—for which data have not been observed—is the “problem at hand.”
- We assume that previous observations of other problems have generated data that can be used in validation assessments.

The committee identified several over-arching principles of VVUQ.

- VVUQ tasks are inter-related
  - Solution verification and propagation of input uncertainties should be integral parts of any validation assessment, for example.
- VVUQ should be applied in the context of specified Quantities of Interest (QOIs).
  - If not, VVUQ questions are not well posed.
- Verification and Validation are not yes/no questions with yes/no answers.
  - Solution verification attempts to quantify or bound numerical error.
  - Validation attempts to quantify or bound model error.

Verification Principles and Practices

Principle: Verification is best performed on software created under appropriate software-quality practices.

Best practice:

- **Use software configuration management and regression testing.**
- **Strive to understand and improve the degree of code coverage attained by regression suites.**
- **Code-to-code comparisons can help, especially in early stages of development, but they do not by themselves constitute sufficient code or solution verification.**
- **Compare against analytic solutions, including those generated by the method of manufactured solutions.**

Principle: Solution verification is well defined only in terms of specified QOIs.

Best practices:

- **Clearly define QOIs for a given solution-verification assessment.**
  - **Different QOIs will be affected differently by numerical errors.**
- **Solution verification should encompass the full range of inputs that will be employed during UQ assessments.**
  - **Numerical error may differ for different values of input parameters.**

Example QOIs:

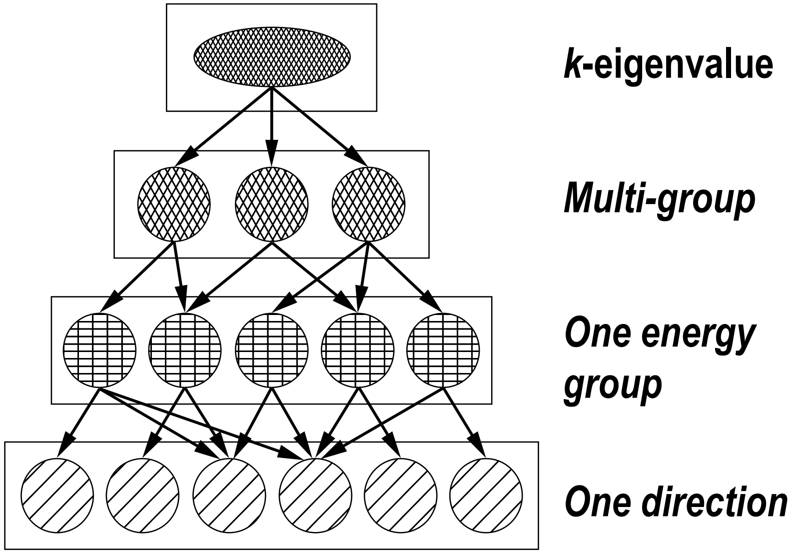
Peak power density in a proposed nuclear reactor.

Load under which a proposed bridge will fail.

Principle: Code and solution verification can be enhanced by exploiting hierarchical compositions.

Best practices:

- Identify hierarchies in mathematical models.
- Design codes with hierarchical code verification in mind.
- Begin code and solution verification at lowest levels of hierarchy, then move upward.



Example:
neutron transport in reactor

Principle: Solution verification should estimate numerical error *for the problem at hand.*

Best practices:

- **When possible, use goal-oriented *a posteriori* error estimates (which estimate numerical error for specified QOIs in the problem at hand).**
- **If goal-oriented *a posteriori* estimates are not available, use self-convergence studies for the problem at hand, if possible.**
- **If possible, control numerical error so that the uncertainty it causes is smaller than those from other sources.**

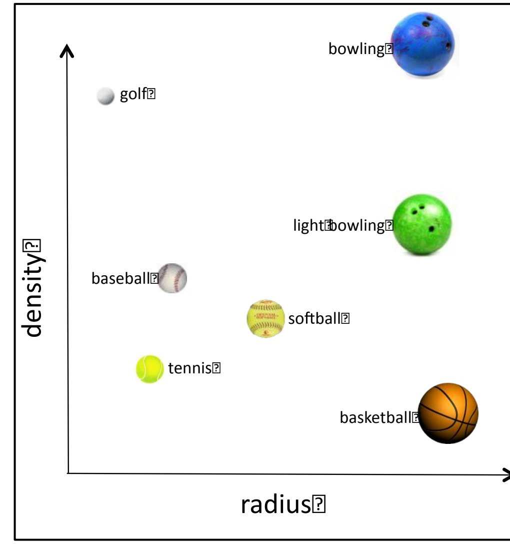
Remark: for many problems of interest in science and engineering, these practices are not possible today.

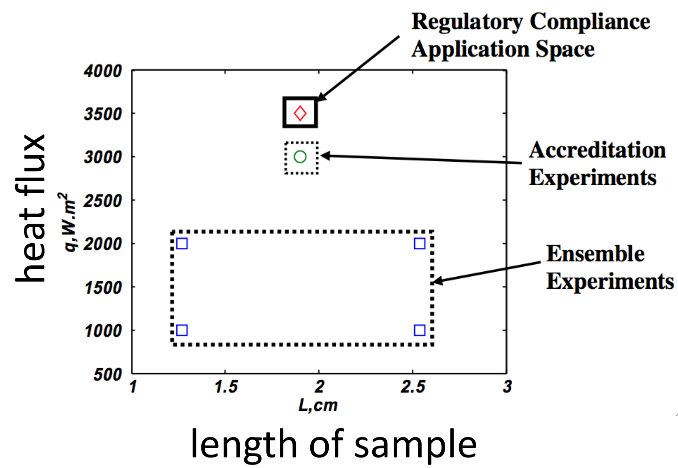
  - ***R&D should broaden availability of 1st practice, which helps enable the 3rd.***
  - ***Better algorithms and computers can broaden availability of the 2nd.***

Validation Principles and Practices

We speak in terms of a "domain of applicability."

- **This appealing concept is useful in validation and prediction.**
  - It helps in assessing relevance of validation data to the problem at hand.
- **Problem features/descriptors form axes that define a "domain space."**
  - Problems map to points in the space.
  - If a new problem is "surrounded" by validation problems, relevance appears high.
- **BUT: DoA relies on judgment (not math).**
  - Who chooses the axes? Omission of an important axis could be fatal.
  - What if the new problem is not "surrounded" by validation problems—the usual case given lots of axes? How do we assess relevance & quantify any added uncertainty?





Principle: A validation assessment informs about model accuracy only in the "domain of applicability" covered by its physical observations.

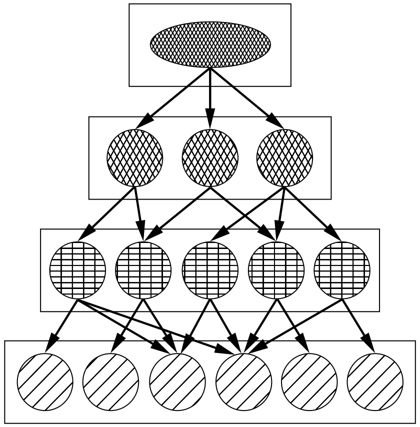
Best practices:

- **Given a QOI and a problem at hand, assess relevance of supporting validation assessments.**
  - **Validation assessment used data from different problems/experiments, often with different QOIs.**
  - **Subject-matter expertise must inform the assessment of relevance.**
- **Use "holdout tests" to test validation and prediction methodologies.**
  - **If methodology does not "predict" a held-out validation dataset, there is little justification for believing the prediction of the problem at hand.**

Principle: Validation assessments can be improved by exploiting hierarchical composition of models.

Best practices:

- **Identify hierarchies in mathematical and computational models.**
- **Seek physical observations that facilitate hierarchical validation assessments.**
- **If possible, use physical observations to constrain uncertainties in model inputs and parameters.**
  - **This is "calibration."**
  - **This is best done at lowest levels of hierarchy, where causes and effects are clearer.**



Principle: Validation assessments must account for errors and uncertainties in physical observations.

Best practices:

- **Identify all important sources of uncertainty and error in the measured data used for validation. Quantify the impact of each on the inferred QOI.**
  - **Sources include instrument calibration, uncontrolled variation in initial/boundary conditions, variability in measurement setup, random variations in physical processes, etc.**
- **Use replicates to inform about variability and measurement uncertainty.**

Remark: assessing measurement uncertainties and errors is often complicated by the fact that the "measured" QOI is actually inferred from measurement of something else.

Prediction Principles and Practices

Principle: Uncertainty in prediction of a QOI must be aggregated from uncertainties from many sources.

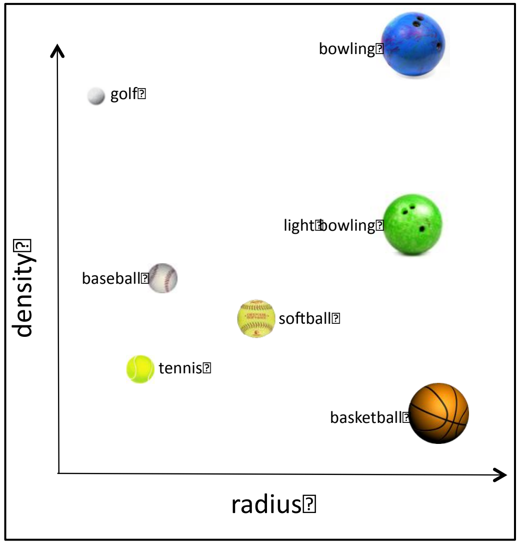
Sources include *model discrepancy*, *numerical errors*, *code errors*, and *uncertain values of input parameters*.

Best practice:

- **Assess sensitivity of the predicted QOI, and of its assessed uncertainties, to each important source of uncertainty and to key assumptions and omissions.**
- **Document key judgments—including those regarding relevance of validation studies to the problem at hand—and assess sensitivity of the QOI (and its uncertainties) to reasonable variations in these judgments.**

Closing remarks

Observation: Judgment informed by subject-matter expertise plays a substantial role in predictions.



- Who chose radius and density to define the ball-drop domain space? Why?
- Do other features matter?
  - Temperature, pressure, humidity, wind, height, surface roughness, elasticity, …
- Are validation data from tennis ball, golf ball, and basketball relevant for the softball?
- How do we map validation-data model error to problem-at-hand model error? No math prescription will work in general.
- Peer review may buy some insurance.
- We do not see anything foolproof.

Remark: It is premature to specify best methodologies for VVUQ tasks.

- We have identified best practices, but we deliberately do not identify best methodologies.
  - Example: We identify that a best practice is to assess sensitivity of a QOI to each source of uncertainty. We do not specify a method for quantifying this sensitivity.
- Method development and improvement are active research areas—the field is in flux.
- Today, some methods work better for some applications while others are better for other applications.

Still to come from the committee …

- **Wei Chen: Educational changes to foster advances in VVUQ methods and applications.**
- **Omar Ghattas: Research needed to improve mathematical foundations of VVUQ.**