Verification and Validation

ESD.864
Noelle Selin

February 12, 2013

Massachusetts Institute of Technology
Engineering Systems Division

Modeling: different depending on decision-making and policy

- How to model for science
- Conducting validation and verification
- Understanding large-scale, complex physical processes requires increasingly elaborate modeling, with associated validation, verification, truth issues (Oreskes)

- All models are wrong, some models are useful.

Types of models

- Simple box models
- Optimization models
- System dynamics
- Conceptual models/frameworks
- Natural system models
- Integrated assessment models
- Life-cycle analysis models
- Mental models

Museum of Science Visit: Many more! …Problem Set #1

Policy Applications for Models

- Agenda-setting
- Screening and assessment
- Scenario evaluation
- Improving understanding
- Guiding negotiations
- Decision-making tools
- Informing the public

Meta-issues

- Scale
- Uncertainty
- Boundaries
- Complexity
- Transparency
- Normative considerations

Models, Verification, Validation

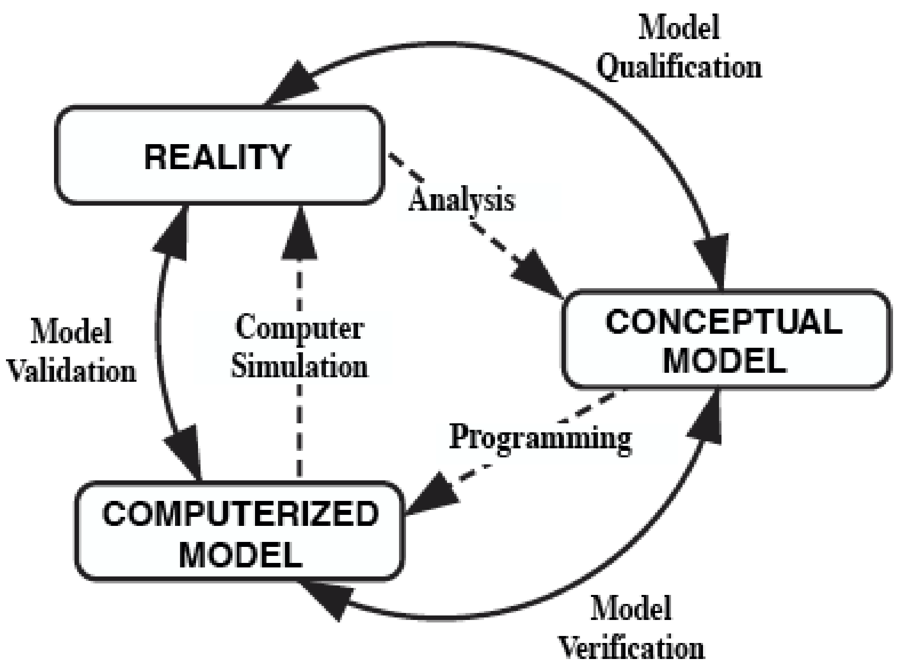


Figure 1
Phases of Modeling and Simulation and the Role of V&V [297]

Source: Oberkampf, William L., and Timothy G. Trucano. "Verification and Validation in Computational Fluid Dynamics." *Progress in Aerospace Sciences* 38, no. 3 (2002): 209-72.

The Modeling Process

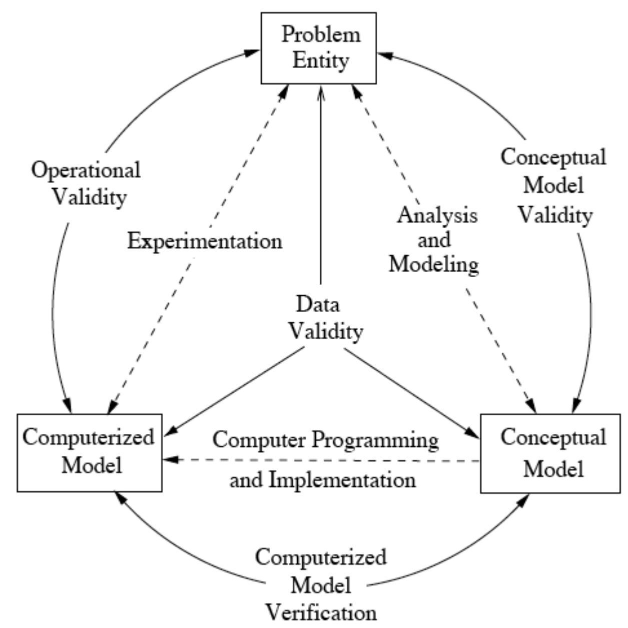


Figure 2: Simplified Version of the Modeling Process

Source: Sargent, Robert G. "Verification and Validation of Simulation Models." In *Proceedings of the 30th conference on Winter Simulation*, pp. 121-30. IEEE Computer Society Press, 1998.

From Sargent, 1998

Steps in model formulation (a science perspective)

- 1. Defining the purpose of the model
- 2. Determining scales of interest
- 3. Determining the dimension of the model
- 4. Selecting processes
- 5. Selecting variables
- 6. Selecting a computer architecture
- 7. Coding the model
- 8. Optimizing the model
- 9. Time steps and intervals
- 10. Initial conditions
- 11. Boundary conditions
- 12. Input data
- 13. Ambient data
- 14. Interpolating data and model results
- 15. Statistics and graphics
- 16. Simulations
- 17. Sensitivity tests
- 18. Improving the model

Design

Logistics

Inputs and data

Analysis and evaluation

Iteration and feedback

[from Jacobson, Fundamentals in Atmospheric Modeling]

Model Design

- 1. Defining the purpose of the model
  - Need to take into account users and their views
  - Understanding goals of the exercise
- 2. Determining scales of interest
  - Match between policy scale and physical scale
- 3. Determining the dimension of the model
  - How much complexity is necessary for science? For policy?
- 4. Selecting processes
  - What processes are decision-relevant?
- 5. Selecting variables
  - Which variables are potentially controllable and which are not?

Design

- "A model should be developed for a specific purpose (or application) and its validity determined with respect to that purpose. If the purpose of a model is to answer a variety of questions, the validity of the model needs to be determined with respect to each question." (Sargent, 1998)

Logistics

- 6. Selecting a computer architecture
  - Will others be able to use the model easily?
  - Excel vs. Matlab vs. Fortran….
- 7. Coding the model
  - Transparency becomes important
- 8. Optimizing the model
  - Jacobson means computationally…

Inputs and Data

- 9. Time steps and intervals
  - Matching the temporal and spatial scale to appropriate decision-making scales
- 10. Initial conditions
  - Data constraints and practical considerations for data selection
- 11. Boundary conditions
  - Defining what's "in" and what's "out" is in practice a science-policy negotiation
- 12. Input data
  - Is itself a model construction
- 13. Ambient data

Analysis and evaluation

- 14. Interpolating data and model results
  - Matching scale of analysis with scale of decision-making
- 15. Statistics and graphics
  - Comprehension and communication
- 16. Simulations
  - Scenarios, hypotheticals, policy options?
- 17. Sensitivity tests
  - How to consider uncertainty, and how to present it? How will uncertainties be understood?

Iteration and Feedback

- 18. Improving the Model
  - One of most crucial steps

Validation and Verification

- Verification: characterization of numerical approximation errors associated with a simulation
  - Did you build the model right?
- Validation: assessment of model accuracy by way of comparison of simulation results with experimental measurements
  - Did you build the right model?
- Uncertainty quantification

See: Roy and Oberkampf, 2011; Sargent, 1998

Methods for V&V

- V&V as part of model development process
- Independent Verification & Validation
  - Performed by a third party
  - Can include accreditation
  - Can be costly, time consuming
- Scoring model according to performance on category scales
  - More on this approach when we talk about NASA credibility scale

Standards

- Degree of V&V required can vary by model type
- IEEE standard for V&V processes:

V&V processes consist of the Verification Process and Validation Process. The Verification Process provides objective evidence for whether the products perform the following:

a) Conform to requirements (e.g., for correctness, completeness, consistency, and accuracy) for all activities during each life cycle process

b) Satisfy the standards, practices, and conventions during life cycle processes

c) Successfully complete each life cycle activity and satisfy all the criteria for initiating succeeding life cycle activities (i.e., builds the product correctly)

The Validation Process provides evidence for whether the products perform the following:

— Satisfy system requirements allocated to the products at the end of each life cycle activity

— Solve the right problem (e.g., correctly model physical laws, implement business rules, and use the proper system assumptions)

— Satisfy intended use and user needs in the operational environment (i.e., builds the correct product)

Source: "IEEE Standard for System and Software Verification and Validation."

Verification

- “ensuring that the computer program of the computerized model and its implementation are correct” (Sargent, 1998)
- Potential verification tasks:
  - Trace intermediate simulation output
  - Test with a known/simplified case and statistically compare results
  - Use animations
  - Structured walk-throughs
  - Dynamic testing (incl. extreme values)
  - Reproducibility

Validation

- “substantiation that a computerized model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model” (Schlesinger et al. 1979)
- Possible validation tasks
  - Uses real-world experimental data
  - Sensitivity analysis
  - Graphs and confidence intervals

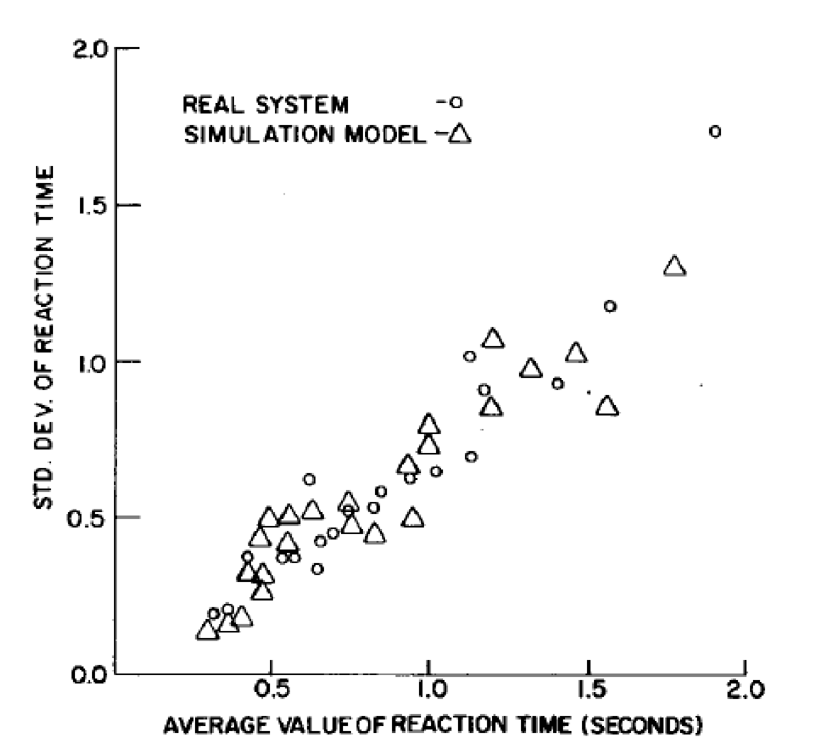
Approaches to validation

	OBSERVABLE SYSTEM	NON-OBSERVABLE SYSTEM
SUBJECTIVE APPROACH	• COMPARISON USING GRAPHICAL DISPLAYS • EXPLORE MODEL BEHAVIOR	• EXPLORE MODEL BEHAVIOR • COMPARISON TO OTHER MODELS
OBJECTIVE APPROACH	• COMPARISON USING STATISTICAL TESTS AND PROCEDURES	• COMPARISON TO OTHER MODELS USING STATISTICAL TESTS AND PROCEDURES

Source: Sargent, Robert G. "Verification and Validation of Simulation Models." *In Proceedings of the 30th conference on Winter Simulation*, pp. 121-30. IEEE Computer Society Press, 1998.

Sargent, 1998

Possible validation output



- Key concept: data can support or refute hypotheses of validity, non-validity for a particular purpose

Source: Sargent, Robert G. "Verification and Validation of Simulation Models." In *Proceedings of the 30th conference on Winter Simulation*, pp. 121-30. IEEE Computer Society Press, 1998.

Sargent, 1998

Uncertainties

- Aleatory: inherent variation in a quantity that can be quantified via a probability density distribution
- Epistemic: uncertainty due to lack of knowledge
- Sometimes combined

Uncertainty quantification

- Roy and Oberkampf outline a method for VV&UQ that propagates relevant uncertainties
- Will talk even more about this in uncertainty unit!

Oreskes: Verification, validation, and Confirmation of numerical models

- Numerical models can’t be verified (by definition), because they aren’t closed systems
  - Input parameters aren’t completely known
  - Continuum mechanics: loss of info at scale < averaging
  - Observation and measurement of variables have inferences, assumptions
  - Nonuniqueness/underdetermination

Validation

- =legitimacy (keep this in mind next class); hasn't been nullified
- Used interchangeably with verification (wrong!), or to imply that model is accurate representation of reality (wrong again!)
- Actually, "validation" as commonly used demonstrates consistency within system or between systems

Numerical models

- “Verification” of numerical solutions: doesn’t imply realism (“bench-marking”: reference to known standard)
- Earth sciences inverse problem: dependent known, independent unknown
  - “Calibration”, “Tuning”, “Empirical adequacy”
- Past history is no guarantee of future performance!

Confirmation

- Fallacy of "affirming the consequent"
- Numerical models are highly complex scientific hypothesis, which cannot ever be certain
- Must use other kinds of observations; verification impossible!
- Confirmation can support probability of model being true

Policy: what now (Oreskes)?

- Need neutral language for evaluating model performance (relative, not absolute)
- Primary value of models is heuristic
- Models are most useful to challenge existing formulations, rather than validate/verify

Responses: Sterman, Rykiel

- Sterman: not limited to earth sciences/complex models! Responsibility of model *consumers*
- Rykiel: definitions and semantics?
- Discussion question: Do you think it's just semantics or is the distinction between verification, validation and confirmation helpful?

More Discussion Questions

- How does model formulation change when considering policy application?
- How does policy analysis change when considering model results?

Problem Set #1

- Reflect on understanding of models by non-technical experts
- Practice V&V on a simple model
- Identify V&V in your own experience
  - And share with others: examples of V&V?

ESD.864 / 12.844J Modeling and Assessment for Policy
Spring 2013