

Lecture #12 - Model evaluation and validation

The fundamental question of model validation is what makes you as the modeler, and your audience, think the model is “right.”

“Right” about what? In the end, right about policy implications and decisions that people want to take supported by insights from the model and the modeling process.

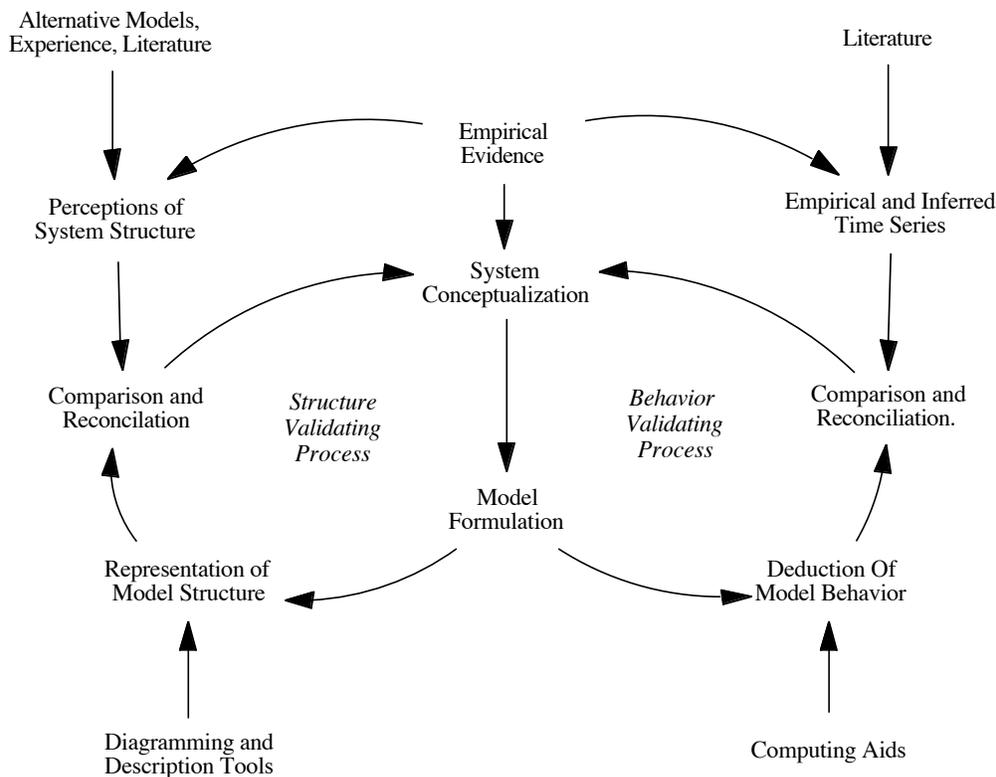
But to get to that end, where the focus is on confidence for policy implications and decisions, the modeler and the audience must focus on a large number of intervening issues, which in combination lead to overall confidence in the work.

Confidence in the overall effort ranges from the details of individual equations to the adequacy of the system boundary selected. People who look at model equations will lose confidence if the algebra looks arcane or contrived, if the form of equations appears strange or unusual, if they contain unexplained or unnamed parameters, if the units don’t check, and so on. Model validation is thus founded on apparently correct equations with a clear algebraic structure that communicates rather than obfuscates.

We talk about “confidence” and model “evaluation” rather than “validation” because no model has ever been “validated” (proved true) and none ever will. We strive for models that are *suitable* for the problem and the purposes of the study and that are *consistent* with all that is known about the slice of reality the model addresses. And we strive for models that are *useful*.

Much of our confidence comes from our careful modeling *process*.

Saeed’s view of the modeling process:



Forrester's tests of validity (Forrester & Senge, Building confidence in system dynamics models, *TIMS Studies in the Management Sciences* 14 (1980):209-228.)

[Main tests indicated by •]

- Tests of model structure
 - Structure-verification test
 - Parameter-verification test
 - Extreme-conditions test
 - Boundary-adequacy test for structure
 - Dimensional-consistency test
- Tests of model behavior
 - Behavior-reproduction tests
 - symptom generation
 - frequency generation
 - relative phasing
 - multiple modes
 - behavior characteristics
 - Behavior-prediction tests
 - pattern prediction
 - event prediction
 - changed-behavior prediction (see below)
 - Behavior-anomaly test
 - Family-member test
 - Surprise-behavior test
 - Extreme-policy test
 - Boundary-adequacy test for behavior
 - Behavior-sensitivity test
- Tests of policy implications
 - System-improvement test
 - Changed-behavior-prediction test
 - Boundary-adequacy test for policy
 - Policy-sensitivity test