

# I Don't Wanna Grow Up...Stuck at Predictive Capability Maturity Model Level Zero

S.M. Kelly, W.J. Rider, R.F. Barrett and S.D. Hammond  
Sandia National Laboratories  
Albuquerque, New Mexico, USA

## Abstract

Under the DOE/NNSAs Advanced Simulation and Computing (ASC) program, Sandia Labs has been developing the Predictive Capability Maturity Model (PCMM) [2], [3] as a means of assessing the completeness (i.e., maturity) of, as well as facilitating project planning and communication for, its computational modeling and simulation activities. The PCMM establishes criteria for maturity levels 0, where very little verification or validation is performed, through 3, where detailed activities are undertaken, and can be implemented through an inclusive process assisting planning and communication. The criterion addresses elements such as fidelity, verification, validation, and uncertainly quantification (UQ). The rigor associated with higher maturity levels is appropriate as simulation results are used as part of higher consequence-simulation-informed decisions. It is the position of the authors that with few exceptions current HPC computer hardware models and simulators cannot defend their predictive capabilities. Given the potential consequences of erroneous results in mis-directing Exascale hardware and software development opportunities, it is our opinion that, at a minimum, simulator and model developers need to thoroughly and properly address model validation. In this position paper we briefly introduce PCMM and appeal for consensus that predictions from simulators or performance model be more thoroughly evaluated.

## I. NOTHIN' EVER SEEMS TO TURN OUT RIGHT

Verification & Validation (V&V), UQ, and/or PCMM do not replace the necessity of relying upon expert judgment, but rather assist in putting its role into the proper context. A large breadth of skills and perspective is necessary to conduct predictive modeling. These range from theoretical modeling to the operational experience on leading edge hardware. The set of activities structured within a V&V plan provide a foundation to communicate the nature of credibility for modeling results. This credibility assessment provides decision makers with the necessary context to properly judge the role modeling plays in a decision. Without V&V the decision makers rely on their faith in the expertise of the modeler or the simulation tools being used.

## II. HOW DO YOU MOVE IN A WORLD OF FOG?

A typical modeling team has the tendency to focus on one aspect of the modeling activity associated with their area of expertise. An example would be a highly calibrated model for a communication subsystem. The calibration would provide a certain degree of modeling fidelity particularly for the systems that are similar to the calibration data source. This team might ignore other sources of data, analytical models, or the sensitivity or uncertainty in the modeling. PCMM, if properly employed, would not allow for this. PCMM provides a structure where the entirety of the modeling activity is highlighted, and each of the necessary elements is examined. By applying PCMM to the modeling activity, each of the technical areas is examined, graded and evidence is collected. PCMM will not perform the work for the user, but it will inform all the stakeholders about what work should and has been done, and asks the stakeholders to consider how well it needs to be done.

## III. I DON'T WANNA HAVE THE BIGGEST AMOUNT

Key to the PCMM is the collection and dissemination of evidence of modeling quality. This includes software quality, code verification (comparison with analytical solutions), model fidelity, validation (comparison with observations), and uncertainty/sensitivity analysis. Any modeling study should be able to provide artifacts of each of these activities to support the assessment of the quality of work. In this way the PCMM provides a method of structuring the support for claims made. In this manner the decision maker has full knowledge and greater transparency when weighing the evidence used in making a decision.

Headings are derived from the lyrics of Tom Waits' song, which are property and copyright of the owners.

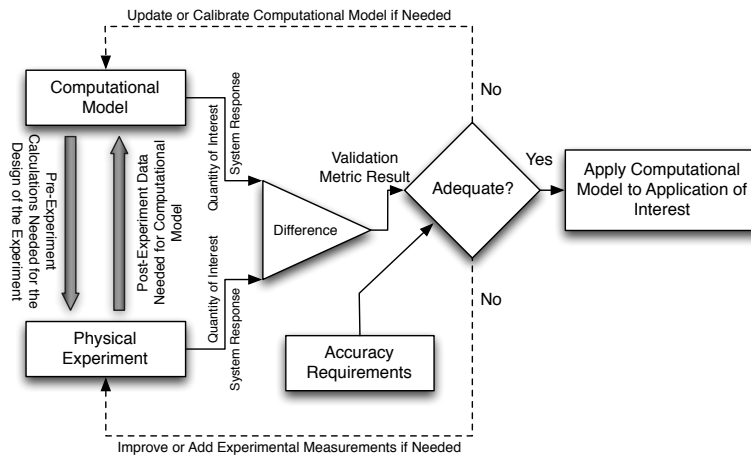


Fig. 1. Validation, Calibration and Prediction [1]

#### IV. I DON'T WANNA BE FILLED WITH DOUBT

When any modeling prediction is made, there is an element of uncertainty in the result. The magnitude of the uncertainty is an important metric in weighing the credibility of a model. Uncertainty quantification is a seminal scientific endeavor and helps to provide the proper context for a result whether it is observational or modeling based. Uncertainty estimation provides an essential sanity check for the modeling and provides the consumer of the results with the understanding of the range of possibilities for the real systems of interest. Ultimately, these uncertainties will be examined against reality in the process of validation.

#### V. WHEN I SEE THE PRICE THAT YOU PAY

Validating a prediction is not a straightforward linear process where a recipe is applied. Rather it is at its core a fundamental application of the scientific method. Given a model for reality, validation is the examination of the model's view of reality against the observations of the real system. The goal is to understand the real system. The fundamental belief that modeling is grounded upon is that a predictive model implies a detailed understanding of the true mechanisms determining the behavior of the real system. Validation is the test of this understanding and offers evidence of its presence.

#### VI. ALL RIGHT THEN, SOUP FROM A STONE?

The physical-modeling based V&V, UQ and PCMM processes have been applied to large scale engineering projects. We are proposing using the same rubric to provide structure to the planning, communication and assessment of this large scale R&D activity. No it's not easy and it's never fully finished. But what ingredients can we collect from CPMM and possibly other predictive assessment tools? Can we find a reasonable, hopefully adoptable, methodology that will apply to sets of hardware models? Is the flowchart in Figure 1 (reproduced from [1]) something we can embrace? Can we specify a criterion or common understanding for how not to cheat in calibration? Or should we apply pixie dust and never grow up ala Peter Pan?

#### ACKNOWLEDGMENTS

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

#### REFERENCES

- [1] W.L. Oberkampf and M.F. Barone. Measures of Agreement between Computation and Experiment: Validation Metrics. Technical Report SAND2005-4302, Sandia National Laboratories, Albuquerque, NM, 2005.
- [2] W.L. Oberkampf and C.J. Roy. *Verification and Validation in Scientific Computing*. Cambridge University Press., 2010.
- [3] W.L. Oberkampf, M. Pilch, and T.G. Trucano. Predictive Capability Maturity Model for Computational Modeling and Simulation. Technical Report SAND2007-5948, Sandia National Laboratories, Albuquerque, NM, 2007.