Validation: Definitions or Descriptions?

Patrick J. Roache
Consultant
1215 Apache Drive
Socorro, NM 87801
hermosa@sdc.org

Abstract

The relation of Validation to Code Verification and Calculation (Solution) Verification is re-stated, and the commonly cited definition of validation for CFD code/models is given. Ambiguities and recommended interpretations of this definition are examined

Introduction

The Lisbon III Workshop has been expanded [1] to cover the full range of Verification and Validation. In the present paper based on [2] the commonly cited definition of validation for CFD code/models is given and found to be inadequate. Ambiguities and recommended interpretations of this definition are examined.

The Scope of V&V

As noted elsewhere [3], "Verification and Validation" or V&V really consists of three subjects, not just two. Code Verification consists of demonstrating that the code is correct; it is capable of achieving correct mathematical solutions to the governing continuum equations in the limit of $\Delta \rightarrow 0$, and the order of convergence is verified at least for well-behaved problems. Arguably the most convincing approach to Code Verification uses grid (or mesh) convergence testing on a nonlinear problem with an exact closed-form solution established by the Method of Manufactured Solutions (MMS) [3]. Although valuable and necessary, Code Verification says nothing about the accuracy of any grid discretization for some new problem.

That accuracy is established during Calculation (or Solution) Verification, in which discretization errors and numerical uncertainties are estimated for a particular problem, without knowledge of the exact solution. The most frequently used method for numerical uncertainty estimation, endorsed by the *ASME Journal of Fluids Engineering* [4], is the Grid Convergence Index (GCI) [3] and its more robust Least-Squares variants [5]. The asymptotic ordered error estimates provided by Richardson Extrapolation are multiplied by a Factor of Safety Fs that has been empirically determined to provide ~ 95% coverage, i.e. an "extended" uncertainty estimator $U_{95\%}$. Single grid error estimators such as Zhu-Zienkiewicz (ZZ) can also be used to obtain uncertainty estimates (for their computed error quantities, which may or may not be of engineering interest) using the same Factor of Safety Fs, at least asymptotically, since all ordered error estimators will agree asymptotically. Calculation Verification is meaningless unless preceded by Code Verification, since coding errors will not necessarily be detected during grid convergence tests of a real problem (with unknown solution).

So Code Verification by itself is not sufficient, nor is Calculation Verification by itself. Together, they establish only mathematical correctness and accuracy, which says nothing about physical accuracy. That is, a meaningless continuum model (e.g. for turbulence) could be correctly implemented in a code and demonstrated to converge as Δ^p where p = 1 or 2 or other, and the numerical error and uncertainty $U_{95\%}$ could be well estimated and small, yet the correspondence to physical reality could be poor. The latter evaluation is the work of Validation.

Validation Definition

Validation: The process of determining the degree to which a model {and its associated data} is an accurate representation of the real world from the perspective of the intended uses of the model.

This definition is widely accepted and cited in CFD, and probably appears acceptable on first reading. In fact, it is inadequate, and disagreement exists on what this definition means or should mean.

There are at least three contested issues.

(1) Does *degree* imply acceptability criteria (pass/fail)?
(2) Does *real world* imply experimental data?
(3) Is *intended use* specific or general, or is it needed at all?

This gives $2^3 = 8$ possible interpretations of the same definition (without even getting into arguments about what is meant by *model*). For a brief history of the definition, see [2]. The job of sorting out claims and arguments is further complicated by the fact that participants in the debates (myself included) have sometimes switched sides on one or more of these issues.

Issue #1. Acceptability Criteria

Concerning the first issue of whether acceptability criteria (or pass/fail criteria, or adequacy, or tolerance) are included in this definition of validation, initially people generally say "yes" without hesitation. Everyone recognizes that pass/fail decisions must be made in any engineering project. This choice is reinforced by the later phrase "from the perspective of the intended uses of the model" which seems to imply project-specific criteria. But pass/fail criteria are *project* requirements, and do not necessarily need to be included while performing "validation". Rather, one can simply evaluate the agreement between computational results and experimental data (with their respective uncertainties) and present the difference as the ***level*** of validation. This acknowledges that the same validation level (e.g., 5% agreement for base pressure) may be adequate for one application but not for another.

Two distinct processes exist: first, comparison of model predictions with experimental data, leading to an assessment of model accuracy, and second, determination of acceptability of that accuracy level for a particular application. The methods in each process have nothing in common. An extreme view not only includes the pass/fail tolerance in validation, but insists that the acceptable tolerance be specified *a priori*, e.g. see [6,7]. This ties a code/model validation rigidly to a particular engineering project rather than to less specific science-based engineering. (Or worse, it neglects the fact that agreement may be acceptable for one application and not for another.)

It is recommended that criteria for *acceptability* of accuracy are not part of validation. Acceptability of the agreement is part of the next project step, variously called accreditation, certification, or other. However, analysts should be careful to avoid endorsing a code as "validated" if it is clearly unsatisfactory for any reasonable application.

Issue #2. Necessity for Experimental Data

Most engineers and scientists understand "real world" in the definition to imply *real world data*, i.e. what most people would call experimental data. Surprisingly, not everyone agrees, because they want the approval implicit in "validation" without the work of obtaining real experimental data. The recommended interpretation, agreeing with [2,6-9], is uncompromising: no experimental data means no validation.

Issue #3. Intended Use

The requirement for "intended use" sounds good at first, but it fails upon closer thought. Did D. C. Wilcox [10] need to have an "intended use" in mind when he evaluated the k-ω RANS turbulence models for adverse pressure gradient flows? Maybe he had uses in mind, but does a modeler need to have the *same* use in mind two decades later? If not, must the validation comparison be repeated? Certainly not.

It is recommended that specific intended use is not required for validation. A useful validation exercise, such as the subject of this Workshop using the experimental data on turbulent backstep flow of Driver and Seegmiller [11] in the ERCOFTAC database, can be used for code/model validation, with neither the experimenters in 1985 nor modelers in 2008 having a specific use in mind. But it is obvious that experiments designed specifically for a validation exercise with a specific application are more likely to produce data on the relevant metrics with relevant precisions than are experiments designed without applications in mind.

Alternative Description

Instead of agonizing over a rigid definition, we can *describe* validation. The first requirement is the distinction between verifications (mere mathematics) and validation, already noted. Then in general terms, validation involves comparison of modeling results with experimental results. This description has been used in the past, but it is too soft. The trouble is that the difference between model result and experiment is often taken to be the accuracy of the model, when in fact the story is more difficult.

The required improvement is *uncertainty*. We can adequately describe validation as the comparison of model results *and their associated uncertainties* with experimental results *and their associated uncertainties*. A methodology for this comparison (including interpretation of the results) is given in *ASME V&V 20* [8]. It uses accepted, well established quantitative techniques for every aspect of the process, and statistical definitions that are consistent between experimental and modeling methodologies. Whether or not these recommended definitions or descriptions are used, the warning by Tsang [3, pg. 26] still applies: it is meaningless to talk about "validation" without significant further qualifications.

Calibration is not Validation

Calibration is the adjustment or tuning of free parameters in a model to fit the model output with experimental data. It is important to state that calibration is not validation. Calibration is a often a necessary component of (strong sense) model development. But validation occurs only when the previously calibrated model predictions are evaluated against a set of data not used in the tuning [2,6-9]. There is no value in tuning free parameters to obtain a base pressure to match an experimental value, and then claiming code/model validation because the "prediction" agrees with the same experiment. (If all point-values are well matched using a small set of free parameters, this will tend to be convincing in itself, but another data set not used in the tuning would be more so.)

One Final Detail

The validation development in *ASME V&V 20* [8] is given in terms of standard uncertainty u, defined as an estimate of the (sample) standard deviation. The abbreviated version of the method of [8] stated in the announcement for this Workshop [1] is restricted to the commonly used $U_{95\%}$ so that the GCI values can be used directly.

References

1. Eça, L. and Hoekstra, M., 2008, Announcement for the *3rd Workshop on CFD Uncertainty Analysis*, Lisbon, 23-24 October 2008, Instituto Superior Técnico, Lisbon. 1 October 2007.

2. Roache, P. J., 2009, "Perspective: Validation - What Does it Mean?", submitted to *ASME Journal of Fluids Engineering*.

3. Roache, P. J., 1998, *Verification and Validation in Computational Science and Engineering*, Hermosa Publishers, Albuquerque.

4. Celik, I. B., Ghia, U., Roache, P. J., and Freitas, C. J. (2008), "Procedure for Estimation and Reporting of Uncertainty due to Discretization in CFD Applications," *ASME Journal of Fluids Engineering*, Editorial.

5. Eça, L. and Hoekstra, M. (2002), "An Evaluation of Verification Procedures for CFD Applications," 24th Symposium on Naval Hydrodynamics, Fukuoka, Japan, 8-13 July 2002.

6. ASME Committee PTC-60, 2006, *ANSI Standard V&V 10. ASME Guide on Verification and Validation in Computational Solid Mechanics*, 29 December 20.

7. Oberkampf, W. L., Trucano, T. G., and Hirsch, C., 2004, "Verification, Validation, and Predictive Capability in Computational Engineering and Physics," *Applied Mechanics Reviews*, 57(5), pp. 345–384.

8. ASME Committee PTC-61, 2008, *ANSI Standard V&V 20. ASME Guide on Verification and Validation in Computational Fluid Dynamics and Heat Transfer*, (expected) 2008.

9. ASCE/EWRI Task Committee on 3D Free Surface Flow Model Verification/Validation, 2008, *3D Free Surface Flow Model Verification/Validation*, (expected) 2008.

10. Wilcox, D. C., 2006, *Turbulence Modeling for CFD*, DCW Industries, La Canada, CA.

11. Driver, D. M. and Seegmiller, H. L., 1985, "Features of a Reattaching Turbulent Shear Layer in Divergent Channel Flow," *AIAA Journal*, Vol. 23, No. 1, pp. 163-171. ERCOFTAC Classic Database C-30.