

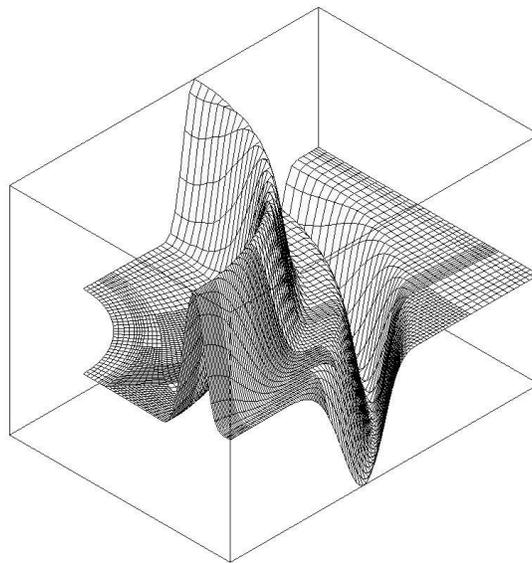
ICES Report 03-52
December 2003

V&V IN COMPUTATIONAL ENGINEERING AND SCIENCE.

PART I: Basic Concepts

by

Ivo Babuska and Tinsley Oden



Institute for Computational Engineering and Sciences (ICES)

The University of Texas at Austin

Report Documentation Page

*Form Approved
OMB No. 0704-0188*

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE DEC 2003	2. REPORT TYPE	3. DATES COVERED -			
4. TITLE AND SUBTITLE V&V in Computational Engineering and Science. Part 1: Basic Concepts		5a. CONTRACT NUMBER			
		5b. GRANT NUMBER			
		5c. PROGRAM ELEMENT NUMBER			
6. AUTHOR(S)		5d. PROJECT NUMBER			
		5e. TASK NUMBER			
		5f. WORK UNIT NUMBER			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Office of Naval Research, One Liberty Center, 875 North Randolph Street Suite 1425, Arlington, VA, 22203-1995		8. PERFORMING ORGANIZATION REPORT NUMBER			
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)			
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)			
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES The original document contains color images.					
14. ABSTRACT see report					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified		15	

V&V IN COMPUTATIONAL ENGINEERING AND SCIENCE.

PART I: Basic Concepts

IVO BABUSKA and TINSLEY ODEN

ICES, The University of Texas at Austin

1. INTRODUCTION

Computational engineering and science, the discipline concerned with the use of computational methods and devices to simulate physical events and engineering systems, is being heralded by many as one of the most important developments in recorded history [10]. Computer predictions of physical events, it is argued, can be of enormous importance in making critical decisions that affect every facet of human existence. As the speed and capacity of computer systems continue to grow, the expectations of users of computer models in decision making continues to grow in kind. Today, some look toward computer-based predictions as a means to obtain vital quantitative information on events that influence the security, health, and well being of much of mankind and many nations and that influence the success of major businesses and other enterprises worldwide.

It is understandable that the reliability of computer predictions has been an issue of growing concern for several decades. Major questions that arise with increasing frequency are: can computer predictions be used as a reliable bases for crucial decisions? How can one assess the accuracy or validity of a computer prediction? What confidence can be assigned to a computer prediction of a complex event? The collection of processes, philosophical issues, and procedures connected with answering these questions has become known as the subject of Verification and Validation (V&V), the verification process addressing the quality of the numerical treatment of the model used in the prediction and the validation process addressing the quality of the model. V&V has been the focus of much study and debate for many years, and a relatively large literature exists on the subject. Among noteworthy works on the subject in the engineering literature are the book of Roache [13], the survey articles of Oberkampf and Trucano and collaborators (eg. [8,9]), the AIAA standards on the subject as it applies to computational

and fluid dynamics [2]. Many other relevant references are cited in these works. Closely akin to works on V&V are those on risk analysis prevalent in applications in, for example, civil engineering [5,6], nuclear engineering [3,12], and in environmental engineering and sciences [1].

The broad interest in V&V in many different scientific and technological areas has led to a diverse and often incompatible list of definitions and concepts as it pertains to different disciplines. Moreover, despite the fact that modern views of the subject have been under development for nearly a decade, much remains to be done toward developing concrete approaches for implementing V&V procedures for particular applications.

In the present paper, we attempt to present a foundation for V&V as it applies to simulations in mechanics and physics and, in a forthcoming second part, to present specific examples of applications of V&V procedures to concrete problems. Our goal is to provide definitions, concepts, and principles that will facilitate communication of ideas and precision of thought regarding the reliability of computer simulations of physical events of interest in computational mechanics and physics. We will then propose specific procedures for implementing V&V processes. We are not concerned with V&V in its broadest (and vaguest) context (modeling manufacturing systems, stock market trends, economics, battle-field models, etc): our interest lies totally in V&V as it applies to the mechanics and physics of material bodies, and generally is motivated by our experience and knowledge of continuum mechanics: solid and fluid mechanics. But we are confident that the framework we develop may apply to a much broader class of disciplines.

2. DEFINITIONS

Here we lay down basic definitions that provide the basis for communicating our views on V&V. Some of these differ from those in standard use. Let us begin with some primitive notions:

Physical Event: An occurrence in nature or in a physical system; a fundamental entity of a physical reality; a physical phenomenon.

The dictionary (Merriam-Webster Collegiate Dictionary, 10th Edition) indicates that an event is “something that happens.” Thus, we are interested in something that happens as a physical reality; not for example, in behavioral aspects or trends in, for instance, sociological or economical systems.

Simulate: To build a likeness; in our case, a likeness produced by an interpretation of output from a computer or computational device.

Mathematical Model (of a physical event): A collection of mathematical constructions that provide abstractions of a physical event consistent with a scientific theory proposed to cover that event.

Data of a Mathematical Model (of a physical event): Factual information that defines the values or ranges of values of parameters in the mathematical model of a physical event.

Discretize: To transform a mathematical model into a finite number of discrete components that can be processed by a digital computer.

Computational Model: The discretized version of a mathematical model that has been designed to be implemented on (or to be processed by) a computer or computational device.

Code: A computer program designed (in the present context) to implement a computational model.

Prediction: (Merriam-Webster Collegiate Dictionary, 10th edition) Something that is predicted, declared or indicated in advance; foretold on the basis of observation, experience, or scientific reason.

Verification: The process of determining if a computational model obtained by discretizing a mathematical model of a physical event and the code implementing the computational model can be used to represent the mathematical model of the event with sufficient accuracy.

Validation: The process of determining if a mathematical model of a physical event represents the actual physical event with sufficient accuracy.

The goal of computer simulation is thus to make predictions of physical events using computational models implemented with appropriate codes. It is important to underscore the intrinsic relationship between a mathematical model of an event and the scientific theory or theories used to characterize the event. As mathematics is, in a fundamental way, the language of science and technology, the mathematical model is to be viewed as a transcription into a concrete and precise format of the theoretical framework in which the modeler intends the event to be depicted. This relationship between the mathematical model and the scientific theory provides the connection between the philosophical underpinnings of V & V and major tenants of contemporary philosophy of science.

If the computational model describes the mathematical model well and the mathematical model relates to the theory well, then the computational model also relates well to the theory.¹

As will become clear later, the models referred to in the definitions of verification and those in validation are, in general, quite different, as are, in some respects, the events they seek to emulate. Also, we emphasize that both verification and validation are *processes*, and the steps in the respective processes may be completely different in that they may pertain to computational and mathematical models different from those selected that the outset to study a particular physical event. Finally, we note that both the verification and validation processes involve determining if the respective processes lead to results of “sufficient accuracy”, leaving open both the meaning of the “sufficient” and of how “accuracy” is to be quantified. We address these issues later in this paper.

3. A VIEW OF V&V

Because many aspects of V&V lie in the intersection of existing theoretical foundations of the mathematical sciences and the reality of events that actually occur in the physical universe, the subject has nourished much discussion along philosophical lines with regard to the place of simulation within the broader boundaries of philosophy of science. Indeed, the notion of validation of a simulation or, more broadly, of a scientific theory, has been the subject of active debate in the philosophy of science for over a half-century.

A first philosophical issue to be confronted concerning V&V is whether or not the validation of models is even possible. The weight of much debate on this subject seems to lie on the negative side: pure and absolute validation is impossible. This point of view is based on the writings of the eminent twentieth-century philosopher, Karl Popper, on the possibility of validating (“verifying” in his words) a scientific theory [11]. Scientific theory, or empirical science as Popper calls it, must be distinguished from logic (or metaphysics); unlike mathematical proofs, which can be established in a logically consistent series of deductions from a set of axioms, scientific theories cannot be proven, they can only be tested through observations. The Popperian argument is that a body of postulates and principles qualifies as a scientific theory only if it is falsifiable; i.e., only if it is possible to find evidence in physical

¹ Reference is made in V&V literature to the need for first constructing a “conceptual model” of an event, referring to the fact that the modeler (or analyst) must go through a mental process of selecting theories or their approximations to represent (to model) an event of interest based on their experience and judgment. While this is undeniably true, the models involved in V & V processes are a step beyond conceptual: they are specific and detailed mathematical models and their discretizations.

observations that contradicts the predictions of the theory. An agreement of observations with predictions does not validate the theory, but once an exception is observed, the theory is judged to be invalid. A theory, can thus, never be validated; it can only be invalidated. In the same spirit, one can argue (as do Konikow and Bredehoeft [4]) that a mathematical model can never be validated; it can only be invalidated.

The same assertion can be applied to verification: neither the computational model of an event nor the code implementing it can ever be completely and absolutely verified; the model and the code can only be judged to be unverified when results are obtained that either show no correlation between solutions of the mathematical model and their computed approximations or that errors exist in the code.

Were this all that could be said of the subject, then the V&V processes could be futile, and there would be no point in going further, or in writing this article. But this is not the case. Popper himself offered hope for a measure of validation of a scientific theory: “So long as a theory withstands severe tests and is not superseded by another theory in the course of scientific progress, we may say that it has ‘proved its mettle’ or that it is *corroborated*.” Thus, while not absolutely validated, sustained success under severe tests can at least corroborate a theory and render it a legitimate basis for decision-making. This notion of corroboration will be adopted as our working view of V&V: *so long as a mathematical model and a computational model of a physical event and the code implementing the computational model withstand detailed and severe tests, we, like Popper, say they have ‘proved their mettle’, and further, that the mathematical model is validated, the computational model is verified, and the code is verified, with respect to a specific series of tests and tolerances.*² Philosophically, *absolute* V&V may be impossible, but V&V relative to a specific series of tests and preset tolerances may be perfectly legitimate as a basis for making decisions.

The V&V processes thus involve the use of measures of accuracy (which is equivalent to the use of measures of error) and specification of tolerances to judge if the accuracy is acceptable. We emphasize that it is not the computational model that one wishes to validate; it is the

² We are very tempted to abandon the term “validation” and use in its place “corroboration”, which, according to Popper, is what is actually done in science: theories are corroborated but never validated. But the long and widespread use of “validation” has imbedded it so deeply in contemporary technical language that we have chosen to continue to use the term but to place on it the burden of many qualifications: validation with respect to specific tolerances for specific quantities of interest. We recognize the danger of putting this more complex notion in the hands of bureaucrats who want to have models and codes absolutely verified and validated.

mathematical model, which corresponds directly or indirectly to a scientific theory or its approximation that covers the physical events of interest.

It is the goal of verification processes to assess the difference between results produced by the computational model and the mathematical model. These types of errors arise in two basic ways³: 1) the code may not be a faithful and accurate implementation of the discretized model and 2) the discretized model may not be an accurate representation of the mathematical model. Thus, verification falls into the two corresponding categories: *code verification*, a province of software engineering, and *solution verification*, which involves a posteriori error estimation. If a code is free of error (an unlikely event), the verification processes are by no means done: the error in the numerical values of the event or events of interest due to discretization remains to be quantified. If this error due to discretization can be quantified, estimated, or bounded in some way and judged to be acceptable by the analyst, then what remains to be assessed is the validity of the theory (i.e. the mathematical model), a goal of the validation process. *Thus, quantifying discretization error, a principal goal of verification processes, is, in general, a necessary prelude to the validation processes, as to do otherwise could lead to an entanglement of modeling and discretization error and obviate the entire validation exercise.*

Code verification involves exercising the computer program developed to implement the computational model to determine and correct coding errors (“bugs”) or other deficiencies that affect the efficiency and quality of output. This is usually done using the code to solve *benchmark problems*; specific, simplified model problems for which accurate solutions or analytical solutions are known. Because benchmark problems rarely involve the data of the model of the actual physical event of interest, they are necessarily incomplete characterizations of the target physical phenomena. Again, the determination of whether a successful solution of a suite of benchmark problems by the code is ample evidence to support its use in the larger calculation is the province of the analyst and is left to the analysts’ judgment (and fallibility) based on his or her experiences and judgment. The use of *manufactured solutions*, which refers to the process of inputting a solution to a model problem and backing out the data needed to produce the solution, yields an exact solution to a problem in the same general class of models for which the predictive calculation is to be performed. Still, the input solution is supplied by the

³ There are, of course, other sources of error occurring in the use of a computerized version of a mathematical model. Round-off error, computer overflow and underflow, etc. can lead to significant errors in computer simulations. These errors are, in general, device-specific. They depend upon the architecture and system software of the computing system(s) on which the code is implemented. These errors can also be detected through tests done with the code run on specific host computers.

analyst and, therefore, its utility as a code verification device will depend upon the skill and insight that the analyst brings to the verification exercise.

Solution verification of a computational model must, in general, be based on a posteriori estimates of the error. By definition, methods of a posteriori error estimation are post-processing techniques that attempt to determine quantitative information on the actual numerical error in computed approximations of the target quantities of interest. Effective methods for such calculations exist and are used with increasing frequency in solution verification processes.

There are important differences and similarities between the verification and validation processes. Validation and verification are similar in that they both involve measuring or estimating errors, selecting tolerances, comparing errors with tolerances, and deciding if the tolerances are met so as to establish a comfortable level of confidence in predictions. Beyond these aspects, the processes are quite different. Verification embodies processes aimed at quantifying the error due to discretization of the mathematical model. In solution verification, the analyst must compare the estimated numerical error in specific quantities of interest to preset tolerances to determine if the computational model is verified to their satisfaction. Solution verification, to a large measure, is thus a mathematical process that can be executed, in theory, to a high degree of precision. While some heuristics are often found in a posteriori error estimation techniques, verification processes exist which are, nevertheless, very effective in accurately quantifying discretization error. The same cannot be said of validation processes; effective processes for comparison of observations with predictions may exist in some applications, but the inductive process of validation (assume test $m+1$ is true if m tests pass the acceptance criteria) is not mathematical induction (if m are assumed true and this implies $m+1$ is true, then all future tests are true). This would be true if models could be absolutely validated, which they cannot.

The models used in predictions and validation may be, and often are, conceptually different. The mathematical model of a physical event may describe a complex system with many thousands of components, each characterized by mathematical submodels with interface conditions and constraints. One discretizes this system model to implement its analysis on a computer. The validation process rarely focuses on the large global model. For example, the verification process may model the behavior of a battleship at high seas; the validation process, in general, cannot, and, so, must be confined to other submodels or component models believed to represent the behavior of key components or processes important in the response of the global model and for which experimental data are available or can be obtained. Each of these

component models becomes a step in both the verification and the validation processes: a verified computational model (in the sense described earlier) of each component must be analyzed and the “verified” results compared with experimental measurements. This process measures the difference between predictions and observations for the sequence of component tests and determines if the error meets tolerances set a priori by the analyst. If the tolerances are not met, the model is invalid. If they are, the analyst must decide, based on judgment, experience, or whim, to accept the prediction of the model based on information obtained in the validation process. All of these issues are in the hands of the analyst (or a panel of peers) and all are based on the judgment, experiences, heuristic arguments, or the evaluations of empirical evidence of the analyst. The global computational (assembled or “system”) model can be verified against specific tolerances, but never absolutely verified. In rare cases in which measurements are made of predicted values in the global model, then the computed values are no longer a prediction, rather they become part of the validation process, termed a *postaudit* [4]. A postaudit enlarges the validation process, and while it may enhance confidence in subsequent prediction, once done, it is no longer a prediction.

Hereafter, we shall refer to the component or subsystem models on which experiments are done in the validation process as *validation problems*.

4. PROCESS AND RULES FOR MODEL SELECTION

The selection of a theory or a number of theories by a computational engineer or scientist (whom we shall continue to call an analyst) to provide the basis of a mathematical model for predicting a physical event is a subjective process that is based on the experience and judgment of the analyst. A mathematical model typically consists of differential, integral, or algebraic equations or inequalities, boundary-and initial-conditions, and other mathematical operations and conditions. Just as the word model means to *imitate* or *emulate*, a mathematical model emulates physical events that are described by scientific theories or approximations of theories. Different analysts may and frequently do develop different mathematical models of the same event. Nevertheless, there are certain rules and procedures that can be followed that can enhance the predictive power of a model. We propose the following rules for model selection and implementation of V&V processes.

Rule 1: Wellposedness and Qualitative Properties. If a mathematical model is to emulate reality, it should possess certain mathematical properties that enable it to produce

qualitatively abstractions of features of the event of interest. These properties can include attributes such as existence of solutions to the governing equations, well posedness, in the sense that solutions exist in some meaningful sense, etc.

Example: If one wishes to model shock interactions in high-speed flows of an inviscid gas over an airfoil, obviously the governing equations of the model must be capable of depicting rapid changes in the momentum field of the flow. If one wishes to model the elastic stability of columns under axial load, the equations modeling the instability must be capable of exhibiting bifurcations in the displacement field.

Rule 2: Quantities of Interest. The particular features of a physical event that are the goal of the prediction should be specified a priori by the analyst. The adequacy of a discretization and a computer implementation of a mathematical model or the adequacy of a mathematical model can be only judged for specific features of the broad event to be simulated. Models may be judged to be adequate for one target output but not another. This rule cannot be overemphasized.

Example: A computer simulation of the flow of air around an airplane may be performed, but the quantity of interest is not, in general, the entire flow field; it may be specifically the drag or the pressure on a specific panel or the temperature on a portion of the leading edge of a wing. Both verification of the computational model and validation of mathematical model must focus on the accuracy with which each of these features is modeled. A model of drag may yield results within preset tolerances when compared to physical observations, but models of other features obtained using the same computational model and mathematical abstractions may be invalid in the sense that they do not fit within the tolerances (see below). The target output or quantities of interest will not, in general, be a single feature of the output, but may involve a substantial list of features.

Rule 3: Select Tolerances. To judge if a computational prediction of a quantity of interest is a sufficiently accurate numerical approximation of the mathematical model of the quantity of interest, quantitative user-specified acceptance criteria must be specified a priori. Usually these criteria are in the form of an acceptable range (tolerance) of the error, i.e. of the difference between the values of the quantities of interest theoretically deliverable by the mathematical model and those obtained from the computational model. Likewise, a mathematical model of an event should be a sufficiently accurate mathematical characterization of a specific feature of a physical event. Quantitative user-supplied acceptance criteria must be specified a priori. The judgment of whether these quantitative tolerances are achieved is based on information obtained

from the validation problems. The appropriate selection of the sequence of validation problems is obviously crucial. Additional analysis of the global problems (leading to the predictions) as, for example, a sensitivity study, could increase the confidence in the estimates of error in the model. Clearly, there is an essential difference between the reliability of the verification of the computation (via a posteriori estimation) and the estimation of the error in the model.

Example: A computational model of the drag developed on the wing of an airplane may be judged to be verified if the computational error in the computed drag is less than 5 percent, or a model of the error in the pressure on a panel of an airfoil is verified if the numerical error is estimated to be less than 8 percent in the L2-norm. The “5 percent” and “8 percent” and the “L2-norm” are user-specified tolerances; different users may have different tolerances for different purposes. The quality of the mathematical model of these quantities of interest is determined on how well the tolerances are met in the validation problems.

Rule 4: Feedback Control. Model selection is, in general, an iterative process. The consequences of selecting an initial model of a physical event can often be assessed with trial calculations obtained using simplified models or by other means. Computed predictions may contradict the original model or may provide results unacceptable to the analyst. This information can be “fed back” and used to make adjustments in the model needed to improve results.

Example: An initial model of the deformation of a structure may be based on linear elasticity theory, but calculations of specific cases may reveal that stresses at certain points in the body exceed the yield point of the materials expected to be used in the application. If these deviate from the linear theory (or if singularities lead to infinite stresses) and these are judged to be important for the particular goal of the simulation, the analyst may determine that plastic deformation of the material should be taken into account. A new model based on elastoplasticity theory can then be constructed and used based on this feedback information and the judgment of the analyst. Again, it is important to note that the acceptability or unacceptability of a model depends upon the particular quantity of interest. The fact that theoretical stresses may be infinite at singular points or surfaces will not necessarily call for a rejection of the model, as these may not affect the values of displacements or strains at points remote from such singularities, which may be the primary targets of the simulations.

Rule 5: Verification Independent of Validation. If a goal of verification and validation is to measure (or estimate) modeling and discretization error, it is necessary to treat these

independently. It is clear that verification of a computational model must be done independently of the validation phase (in general, before validation); else modeling error cannot be distinguished from discretization error. Any validation exercise that is based on a computational model in which discretization error is not quantified is futile, because modeling and approximation error are then intertwined in an indecipherable way.⁴

Rule 6: Data Dependence. Data on material parameters and other features that are used in a mathematical model are obtained from laboratory tests, from various experiments and observations, and other methods. The data, in the present context, includes information on the physical environment of the event to be depicted by the mathematical model; e.g. geometric information, dimensions, boundary data for boundary conditions and initial data for initial conditions, physical coefficients and parameters, data on sources and sinks, etc. The form in which data occurs may be statistical, may include probability density functions, fuzzy sets, etc.⁵ The mathematical model and the data are thus always related. Indeed, the definitions of the mathematical model, in general, include input data, such as coefficients in the governing equations. This rule is that the appropriate mathematical model should be chosen relative to the input data.

Example: If the data are deterministic, the model can be based on deterministic equations. If the analyst is convinced that the fluctuations in data due to uncertainty are small, then the model could utilize perturbation theory. In either case, the choice of the model will take into account the nature of the data.

Rule 7: Convergence. The computational model involves parameters that characterize the discretization of the mathematical model (mesh size, time step, etc.) It should always be theoretically possible to recover the mathematical model and its solutions in the limit as the parameters are appropriately varied. Ideally, this convergence property should be proven mathematically; if this is impossible, convergence should be tested on specific examples.

Rule 8: Reproducibility of Experimental Results. At the heart of the validation process is the concept that experimental results (or observational data) will be obtained in verification problems (the component validation tests) and that these will be compared with those delivered

⁴ We believe that any validation process in which computational parameters, such as mesh size, are varied to bring experimental observations into closer agreement with numerical results are fundamentally flawed. It is our contention that solution verification must be done independently of validation steps. The reverse process is more tenable: experiment observations can be useful in making improvements in mathematical models, but only so long as discretization error is quantified and controlled.

⁵ Frequently, insufficient data is available to make detailed quantitative appraisals. Then, other methods must be employed that do not rely upon complete or extensive information on the properties of the model.

by mathematical models of the events measured (the discretization errors being quantified to the analysts' satisfaction). Unfortunately, the experimental data itself can be, and almost always is, in error. The problem of estimating, controlling, and quantifying experimental error parallels in an astonishing way that of V & V processes for mathematical and computational models. On the one hand, there are the apparatuses needed to make measurements and supply input to a program of physical tests, while on the other hand, there are devices and possibly technicians that can record and interpret the observed output. In analogy to the verification processes, the apparatuses must be calibrated to and lead to accuracies regarded as acceptable by the experimentalist. In the analogy to validation, confidence that the experiment itself measures accurately the event of interest must be established. Obviously, V & V processes may have little value if the range of experimental error is not known and if the experimental results are not *reproducible*.

By *reproducibility of experiments*, we shall mean the variation width in measured values of an event of interest over a set of physical experiments. We will provide a precise mathematical definition of this concept in Part II of this essay. As a general rule, we would hope to apply V & V processes to quantities of interest for which a high level of reproducibility is observed in a series of experiments. Reproducibility will obviously also effect the choice of the mathematical and computational models used in the simulation.

These are basic rules for model selection, mathematical and computational. There are several others that could be listed, but these form the basis for the V&V processes advocated here.

5. SUMMARY

Some essential points presented thus far are summarized as follows:

- Validation involves comparison of observed physical events with those predicted by mathematical models of the component events (validation problems).
- The prediction can be only verified to certain tolerances but not validated.
- By physical event, we mean a specific physical entity, a quantity (or list of quantities) of interest that is (are) generally specified a priori, before the computer simulation.
- The accuracy of a model's prediction can be judged only relative to tolerance supplied by the analyst and then only for a limited number of observations, never for all possible situations

covered by the model; these tolerances are arbitrary, in the sense that different analysts may choose different tolerances for different quantities of interest for different purposes.

- The tolerances for validation are statistical in nature, generally given in terms of a significance level.
- Verification involves two basic components, code verification, which encompasses the software engineering processes of determining if the code faithfully implements the computational model, and solution verification, which is concerned with numerical accuracy with which the mathematical model is approximated by the computational model.
- Solution verification involves assessing the accuracy of computed results for both the global and the validation component models compared with those capable of being predicted by the mathematical models selected to depict the physical event of interest. Solution verification involves a posteriori estimation of a numerical or discretization error.
- Experimental results used in the validation process can themselves be in error. The reproducibility of experimental results for validation component models is essential.

The impossibility of total and absolute verification and validation of computer models, and the dependence of V&V on subjective processes based on human experience, in no way make the subject inferior or different from any other scientific endeavor. Ernest Nagel [7] noted that, “The daily affairs of men are carried out within a framework of steady habits and confident beliefs, on the one hand, and of unpredictable strokes of fortune and precarious judgments on the other.... In spite of such uncertainties, we manage to order our lives with some measure of satisfaction; and we learn, though not always easily, that, even when grounds for our belief are not conclusive, some beliefs can be better grounded than others.” And, finally, “...the methods of the natural sciences are the most reliable instruments men have thus far devised for ascertaining matters of fact, but that withal the conclusions reached by them are only probable because they rest on evidence which is formally incomplete.”

Acknowledgments. We thank a number of colleagues who graciously read early drafts of this paper and who provided constructive advice that led to significant improvements. Among these we mention in particular, Pat Roache, Bill Oberkampf, Hans Mair, Scott Doebeling, Roger Logan, and Charbel Farhat. Any shortcomings, errors, and misconceptions recorded here, however, are our own. We gratefully acknowledge that our work over the last few years on this subject was supported by ONR and NSF grants.

REFERENCES

1. Cullen, A.C. and Frey, H.Ch., *Probabilistic Techniques in Exposure Assessment*, Plenum Press, New York, 1999.
2. *Guide for Verification and Validation of Computational Fluid Dynamics Simulation*, TECH REP, AIAA G-077-1998, American Institute of Aeronautics and Astronautics, 1998.
3. Helton, J.C., "Uncertainty and Sensitivity Analysis in Performance Assessment for WAST Isolation Plant," *Computer Physical Communication* 117, pp.156-180, 1999.
4. Konikow, L. and Bredenhoeft, J.D., "Ground Water Models Cannot Be Validated," *Advances in Water Resources*, Vol. 15, pp. 75-83, and "reply to Comment," Vol. 15, pp. 371-372, 1992.
5. Melchers, R.E., *Structural Reliability Analysis and Predictions*, John Wiley, New York, 1999.
6. Modarres, M., *What Every Engineer Should Know About Reliability and Risk Analysis*, Marcel Dekker, 1993.
7. Nagel, E. "Principles of the Theory of Probability," *International Encyclopedia of United Sciences*, Ed. By O. Newrath, R. Carnap, and C. Morris, The University of Chicago Press, Chicago, Vol. 1, Nos.6-10, pp.341-422, 1955.
8. Oberkampf, W.L., and Truncano, T.G., "Verification and Validation in Computational Fluid Dynamics," *Progress in Aerospace Sciences* Vol.38, pp.209-272, 2002.
9. Oberkampf, W.L., Truncano, T.G., and Hirsch, Ch., "Verification, Validation, and Predictive Capabilities in Computational Engineering, and Physics" *Tech Rep Sand* No. 2003-3769, SANDIA Nat. Lab, 2003.
10. Oden, J.T., "The Promise of Computational Engineering and Science: Will it be Kept?" *Expressions*, Bulletin for the IACM, No. 12, pp. 12-15, 2002.
11. Popper, Karl, *The Logic of Scientific Discovery*, Routledge Classics, Taylor and Francis, London and New York, 2003.
12. Ramsey, Ch. B. and Modarres, M., *Commercial Nuclear Power Assuring Safety for the Future*, John Wiley, New York, 1997.
13. Roache, P.J., *Verification and Validation in Computational Science and Engineering, Science and Engineering*, Hermosa Publishers, Albuquerque, 1998.