ASME V&V 10-2019

[Revision of ASME V&V 10-2006 (R2016)]

Standard for Verification and Validation in Computational Solid Mechanics

AN INTERNATIONAL STANDARD

The American Society of Mechanical Engineers

ASME V&V 10-2019

[Revision of ASME V&V 10-2006 (R2016)]

Standard for Verification and Validation in Computational Solid Mechanics

AN INTERNATIONAL STANDARD

The American Society of Mechanical Engineers

Two Park Avenue • New York, NY • 10016 USA

Date of Issuance: March 6, 2020

The next edition of this Standard is scheduled for publication in 2024.

ASME issues written replies to inquiries concerning interpretations of technical aspects of this Standard. Interpretations are published on the Committee web page and under http://go.asme.org/InterpsDatabase. Periodically certain actions of the ASME V&V Committee may be published as Cases. Cases are published on the ASME website under the V&V Committee Page at http://go.asme.org/VnVcommittee as they are issued.

Errata to codes and standards may be posted on the ASME website under the Committee Pages to provide corrections to incorrectly published items, or to correct typographical or grammatical errors in codes and standards. Such errata shall be used on the date posted.

The V&V Committee Page can be found at http://go.asme.org/VnVcommittee. There is an option available to automatically receive an e-mail notification when errata are posted to a particular code or standard. This option can be found on the appropriate Committee Page after selecting "Errata" in the "Publication Information" section.

This international code or standard was developed under procedures accredited as meeting the criteria for American National Standards and it is an American National Standard. The Standards Committee that approved the code or standard was balanced to assure that individuals from competent and concerned interests have had an opportunity to participate. The proposed code or standard was made available for public review and comment that provides an opportunity for additional public input from industry, academia, regulatory agencies, and the public-at-large.

ASME does not "approve," "rate," or "endorse" any item, construction, proprietary device, or activity.

ASME does not take any position with respect to the validity of any patent rights asserted in connection with any items mentioned in this document, and does not undertake to insure anyone utilizing a standard against liability for infringement of any applicable letters patent, nor assume any such liability. Users of a code or standard are expressly advised that determination of the validity of any such patent rights, and the risk of infringement of such rights, is entirely their own responsibility.

Participation by federal agency representative(s) or person(s) affiliated with industry is not to be interpreted as government or industry endorsement of this code or standard.

ASME accepts responsibility for only those interpretations of this document issued in accordance with the established ASME procedures and policies, which precludes the issuance of interpretations by individuals.

The American Society of Mechanical Engineers
Two Park Avenue, New York, NY 10016-5990

CONTENTS

FOREWORD

Since the mid-1960s, computer simulations have come to dominate engineering mechanics analysis for all but the simplest problems. This reliance on complicated simulations makes a systematic program of verification and validation (V&V) necessary to ensure the accuracy of these simulations. This Standard describes such a program.

The concept of systematic V&V is not new. The software development community has long recognized the need for a software quality assurance (SQA) plan for scientific and engineering products. The Institute of Electrical and Electronic Engineers (IEEE) was the first to publish and adopt guidelines and standards for engineering SQA appropriate for developers. SQA guidelines, while necessary, are not sufficient to cover the issues of computational physics and engineering or the vast array of problems to which end users apply the codes. To fill this gap, the concept of application-specific V&V was developed.

Scientific and engineering communities have been exploring application-specific V&V since the mid-1990s. The Department of Defense's Defense Modeling and Simulation Coordination Office (DMSCO) produced recommended practices suitable for large-scale modeling and simulation in 1996. However, these DMSCO guidelines do not directly focus on the details of computational physics and engineering. The American Institute of Aeronautics and Astronautics produced the first V&V guidelines tailored for detailed analyses in the area of computational fluid dynamics (CFD) in 1998.

Recognizing the need for a similar set of guidelines for computational solid mechanics (CSM), members of the CSM community formed a committee under the auspices of the United States Association for Computational Mechanics in 1999. The American Society of Mechanical Engineers (ASME) Board on Performance Test Codes (PTC) granted the committee official status in 2001 and designated it the PTC 60 Committee on Verification and Validation in Computational Solid Mechanics. In 2008, an overarching committee for multiple V&V application areas was established by ASME as the V&V Standards Committee on Verification and Validation in Computational Modeling and Simulation. ASME reorganized the committees under the V&V Standards Committee, and the PTC 60 Committee was renamed the V&V 10 Subcommittee on Verification and Validation in Computational Solid Mechanics.

The V&V 10 Subcommittee (previously PTC 60 Committee) undertook the task of writing the proposed guidelines. Its membership has consisted of solid mechanics analysts, experimenters, code developers, and managers from industry, government, and academia. Represented industries include aerospace/defense, commercial aviation, automotive, bioengineering, and software development; represented government agencies include the Department of Defense, the Department of Energy, and the Federal Aviation Administration.

Early discussions within the V&V 10 Subcommittee revealed an immediate need for a common language and process definition for V&V appropriate for CSM analysts and their managers and customers. The first edition of ASME V&V 10, Guide for Verification and Validation in Computational Solid Mechanics, described the semantics of V&V and defined the process of performing V&V in a manner that facilitates communication and understanding among the various performers and stakeholders.

The Guide was approved by the V&V 10 Subcommittee and was approved and adopted by the American National Standards Institute in 2006. Since that original edition was released, the issues and problems of V&V in CSM have been studied through discussion and the generation of supporting documentation, including an example problem standard, ASME V&V 10.1. That work contributed to the maturation of the discipline and influenced this revised edition, which is now titled Standard for Verification and Validation in Computational Solid Mechanics.

This Standard is available for public review on a continuing basis. This provides an opportunity for additional public-review input from industry, academia, regulatory agencies, and the public-at-large.

ASME V&V 10 was approved by the V&V Standards Committee on March 28, 2019 and was approved and adopted by the American National Standards Institute on July 23, 2019.

ASME V&V COMMITTEE
Verification and Validation in Computational Modeling and Simulation

(The following is the roster of the Committee at the time of approval of this Standard.)

STANDARDS COMMITTEE OFFICERS

T. M. Morrison, *Chair*
B. H. Thacker, *Vice Chair*
K. M. Hyam, *Secretary*

STANDARDS COMMITTEE PERSONNEL

M. D. Benedict, Air Force Research laboratory
J. Bischoff, Zimmer Biomet
S. W. Doebling, Los Alamos National Laboratory
K. Dowding, Sandia National Laboratories
C. J. Freitas, Southwest Research Institute
Y. Hassan, Texas A&M University
M. Horner, ANSYS, Inc.
K. M. Hyam, The American Society of Mechanical Engineers
J. S. Kaizer, U.S. Nuclear Regulatory Commission
D. M. Moorcroft, Federal Aviation Administration
T. M. Morrison, U.S. Food and Drug Administration
S. Rachuri, U.S. Department of Energy
R. R. Schultz, Consultant
V. Sharma, Exponent, Inc.
B. H. Thacker, Southwest Research Institute
A. F. Emery, *Contributing Member*, University of Washington

V&V 10 SUBCOMMITTEE — VERIFICATION AND VALIDATION IN COMPUTATIONAL SOLID MECHANICS

D. M. Moorcroft, *Chair*, Federal Aviation Agency
J. L. O'Daniel, *Vice Chair*, U.S. Army Corps of Engineers, Engineer Research and Development Center
M. Pagano, *Secretary*, The American Society of Mechanical Engineers
M. C. Anderson, National Nuclear Security Administration
S. Atamturktur, The Pennsylvania State University
M. D. Benedict, Air Force Research Laboratory
J. Budzien, Los Alamos National Laboratory
R. M. Ferencz, Lawrence Livermore National Laboratory
L. Gutkin, Kinectrics
S. L. Kieweg, Sandia National Laboratories
D. Lancaster, Pratt & Whitney
I. Lopez, Lawrence Livermore National Laboratory
W. L. Oberkampf, William L. Oberkampf Consulting
G. Orient, Sandia National Laboratories
C. F. Popelar, Southwest Research Institute
C. Robeck, Thornton Tomasetti - Weidlinger Applied Science
B. H. Thacker, Southwest Research Institute
J. V. Cox, *Alternate*, Sandia National Laboratories
W. R. Witkowski, *Alternate*, Sandia National Laboratories
B. Benowitz, *Contributing Member*, Thornton Tomasetti
S. W. Doebling, *Contributing Member*, Los Alamos National Laboratory
T. K. Hasselman, *Contributing Member*, Consultant
S. R. Hsieh, *Contributing Member*, Lawrence Livermore National Laboratory
X. Jiang, *Contributing Member*, General Electric Co.
R. W. Logan, *Contributing Member*, Consultant
H. U. Mair, *Contributing Member*, Johns Hopkins University Applied Physics Laboratory
T. L. Paez, *Contributing Member*, Thomas Paez Consulting
R. Rebba, *Contributing Member*, General Motors R&D Center
C. P. Rogers, *Contributing Member*, Crea Consultants, Ltd.
J. F. Schultze, *Contributing Member*, Los Alamos National Laboratory
K. Teferra, *Contributing Member*, Consultant

CORRESPONDENCE WITH THE V&V COMMITTEE

General. ASME Standards are developed and maintained with the intent to represent the consensus of concerned interests. As such, users of this Standard may interact with the Committee by requesting interpretations, proposing revisions or a case, and attending Committee meetings. Correspondence should be addressed to:

Secretary, V&V Standards Committee
The American Society of Mechanical Engineers
Two Park Avenue
New York, NY 10016-5990
http://go.asme.org/Inquiry

Proposing Revisions. Revisions are made periodically to the Standard to incorporate changes that appear necessary or desirable, as demonstrated by the experience gained from the application of the Standard. Approved revisions will be published periodically.

The Committee welcomes proposals for revisions to this Standard. Such proposals should be as specific as possible, citing the paragraph number(s), the proposed wording, and a detailed description of the reasons for the proposal, including any pertinent documentation.

Proposing a Case. Cases may be issued to provide alternative rules when justified, to permit early implementation of an approved revision when the need is urgent, or to provide rules not covered by existing provisions. Cases are effective immediately upon ASME approval and shall be posted on the ASME Committee web page.

Requests for Cases shall provide a Statement of Need and Background Information. The request should identify the Standard and the paragraph, figure, or table number(s), and be written as a Question and Reply in the same format as existing Cases. Requests for Cases should also indicate the applicable edition(s) of the Standard to which the proposed Case applies.

Interpretations. Upon request, the V&V Standards Committee will render an interpretation of any requirement of the Standard. Interpretations can only be rendered in response to a written request sent to the Secretary of the V&V Standards Committee.

Requests for interpretation should preferably be submitted through the online Interpretation Submittal Form. The form is accessible at http://go.asme.org/InterpretationRequest. Upon submittal of the form, the Inquirer will receive an automatic e-mail confirming receipt.

If the Inquirer is unable to use the online form, he/she may mail the request to the Secretary of the V&V Standards Committee at the above address. The request for an interpretation should be clear and unambiguous. It is further recommended that the Inquirer submit his/her request in the following format:

Subject:	Cite the applicable paragraph number(s) and the topic of the inquiry in one or two words.
Edition:	Cite the applicable edition of the Standard for which the interpretation is being requested.
Question:	Phrase the question as a request for an interpretation of a specific requirement suitable for general understanding and use, not as a request for an approval of a proprietary design or situation. Please provide a condensed and precise question, composed in such a way that a “yes” or “no” reply is acceptable.
Proposed Reply(ies):	Provide a proposed reply(ies) in the form of “Yes” or “No,” with explanation as needed. If entering replies to more than one question, please number the questions and replies.
Background Information:	Provide the Committee with any background information that will assist the Committee in understanding the inquiry. The Inquirer may also include any plans or drawings that are necessary to explain the question; however, they should not contain proprietary names or information.

Requests that are not in the format described above may be rewritten in the appropriate format by the Committee prior to being answered, which may inadvertently change the intent of the original request.

Moreover, ASME does not act as a consultant for specific engineering problems or for the general application or understanding of the Standard requirements. If, based on the inquiry information submitted, it is the opinion of the Committee that the Inquirer should seek assistance, the inquiry will be returned with the recommendation that such assistance be obtained.

ASME procedures provide for reconsideration of any interpretation when or if additional information that might affect an interpretation is available. Further, persons aggrieved by an interpretation may appeal to the cognizant ASME Committee or Subcommittee. ASME does not "approve," "certify," "rate," or "endorse" any item, construction, proprietary device, or activity.

Attending Committee Meetings. The V&V Standards Committee regularly holds meetings and/or telephone conferences that are open to the public. Persons wishing to attend any meeting and/or telephone conference should contact the Secretary of the V&V Standards Committee. Future Committee meeting dates and locations can be found on the Committee Page at http://go.asme.org/VnVcommittee.

PREFACE

The ASME V&V 10 Subcommittee on Verification and Validation (V&V) in Computational Solid Mechanics is creating a family of standards that together present a comprehensive picture of the standards and practices governing this process.

(a) ASME V&V 10-2006, Guide for Verification and Validation in Computational Solid Mechanics, was the first edition of ASME V&V 10. Intended as an overview of V&V, it also included background material and definitions necessary to understand the other standards in the series. It contains definitions of key terms associated with V&V, and it provides context for the role of V&V in engineering as well as an overview of key aspects of application. ASME V&V 10-2019 is the first revision of that Guide. Since publication of the first edition, the field of V&V has matured to the point that ASME V&V 10's title has been changed from "Guide" to "Standard."

(b) ASME V&V 10.1-2012, An Illustration of the Concepts of Verification and Validation in Computational Solid Mechanics, is a follow-on publication that illustrates the steps in the V&V process described in ASME V&V 10-2019 through a worked example. This Standard is intended to provide a more concrete look at how to translate the process of V&V 10-2019 into the reality of an engineering project.

(c) ASME V&V 10.2 is currently under development with the working title The Role of Uncertainty Quantification in Verification and Validation of Computational Solid Mechanics. This Standard is intended to take a deeper look at the importance of uncertainty quantification (UQ), types and characterization of uncertainties, introduction to UQ methodologies, and how UQ is applied during each phase of the V&V process.

(d) ASME V&V 10.3 is currently under development with the working title The Role of Validation Metrics in Verification and Validation of Computational Solid Mechanics. This Standard is intended to provide a primer of mathematical metrics to measure the difference between calculated results and either analytical or semianalytical solutions (in the case of code verification) or experimental measurements (in the case of validation).

Readers are encouraged to begin with ASME V&V 10-2019 as it lays the groundwork, but may find a concurrent reading of ASME V&V 10.1-2012 beneficial, as it closely follows the V&V process described through an example. ASME V&V 10.2 and ASME V&V 10.3 should be read subsequent to ASME V&V 10-2019, as the foundation of ASME V&V 10-2019 is necessary to understand the significance of the deeper treatments in ASME V&V 10.2 and ASME V&V 10.3.

STANDARD FOR VERIFICATION AND VALIDATION IN COMPUTATIONAL SOLID MECHANICS

1 EXECUTIVE SUMMARY

Program managers need assurance that computational models of engineered systems are sufficiently accurate to support programmatic decisions. This Standard provides the technical community — engineers, scientists, and program managers — with guidelines for assessing the credibility of computational solid mechanics (CSM) models.

Verification and validation (V&V) are the processes by which evidence is gathered to determine the accuracy of the computer model for specified conditions. These accuracy results, along with uncertainty quantification (UQ), contribute to the determination of the credibility of the model for the conditions of its intended use.

Professional organizations differ in their definitions of V&V. The American Society of Mechanical Engineers (ASME) V&V 10 Subcommittee on Verification and Validation in Computational Solid Mechanics has chosen definitions consistent with those published by the Department of Defense (DoD) (ref. [1]) and by the American Institute of Aeronautics and Astronautics (AIAA) (ref. [2]). Verification assesses the numerical accuracy of a computational model regardless of the physics being modeled. Both code verification (addressing errors in the software and numerical algorithms) and calculation verification (estimating the numerical errors due to under-resolved discrete representations of the mathematical model) are addressed. Validation assesses the degree to which the computational model is an accurate representation of the physics being modeled. It is based on comparisons between numerical simulations and relevant experimental results. Validation is essential in assessing the predictive capability of the model in the physical realm of interest, and it must address uncertainties that arise from both experimental and computational procedures.

As shown in Figure 2.3-1, the general V&V process begins with a statement of the intended use of the model and pertinent information about the system being modeled so that the relevant physics are included in both the model and the experiments performed to validate the model. Modeling and experimental activities are guided by the response quantities of interest and the accuracy requirements for the intended use. Experimental outputs intended for validation for component-level to system-level tests should, whenever possible, be provided to modelers only after verification and the numerical simulations for those outputs have been performed.

Ideally, the V&V process for a particular application ends with acceptable agreement between model predictions and experimental outputs, after the uncertainties in both have been taken into account. If the agreement between model and experiment is not acceptable, an assessment should be performed to determine why agreement was not met and, potentially, the processes of V&V repeated by updating the model and performing additional experiments. Successful completion of the validation process, demonstrated by satisfactory agreement between simulation and experiment, means that the model adequately reproduces the experimental measurements that have been obtained.

Once the validation process is successfully completed, the model should be assessed to determine if its predictive capability, including relevant uncertainties, is adequate for conditions where no experimental data are available. Since most models are developed for use where experimental data are not available, predictive capability must address a much wider range of uncertainties than validation. This Standard introduces the concept of predictive capability but does not go into detail because of the early stage of development of this field.

Finally, it is important to document all V&V activities. In addition to preserving the compiled evidence of V&V, documentation records the justifications for important decisions such as selecting primary response quantities and setting accuracy requirements. Documentation thereby supports the primary objective of V&V: to build confidence in the predictive capability of computational models.

The guidance provided herein will enable managers and practitioners of V&V to better assess and enhance the credibility of CSM models. Upon reading about the process described in this Standard and illustrated in ASME V&V 10.1-2012, engineers may be left with the sense that the real-world constraints of the engineering environment (i.e., schedule and budget) do not allow for sufficient project scope to complete the V&V process to a satisfactory level of rigor. Users of this Standard are nonetheless encouraged to provide a V&V foundation for their engineering calculations and identify any associated uncertainties and risks.

The ASME V&V 10 Subcommittee recognizes that program needs and resources vary and that the application of the guidance provided herein to specific cases must accommodate budget, schedule, and risk considerations. This Standard explains the principles of V&V so that practitioners can better appreciate and understand how decisions concerning V&V can affect their ability to assess and enhance the credibility of CSM models. It is the assertion of the ASME V&V 10 Subcommittee that some assessment of and improvement to credibility is better than none at all. The time and budget spent on V&V/UQ should be judged based on how much influence modeling and simulation have on the cost, design, safety, and reliability of the system being analyzed, as well as the magnitude of the detrimental consequences resulting from not meeting requirements relative to these aspects.

The V&V 10 Subcommittee advises users of this Standard to exercise caution when using statements such as "This model has been validated" or "This is a validated model." These statements raise questions such as "To what experimental measurements was the model compared? Over what set of conditions? With what model and data uncertainties? To what level of accuracy? Validated to what intended use?" In fact, there is much more value in that set of questions than in the original statements. Perhaps as more and more engineers and stakeholders ask these questions of their own computational results, V&V will become part of standard engineering practice. This will help change the professional culture so that V&V/UQ analyses are integrated into the project planning phase and considered inseparable from modeling and simulation.

2 INTRODUCTION

CSM plays an increasingly important role in the design and performance assessment of engineered systems. Automobiles, aircraft, and weapon systems are examples of engineered systems that have become more reliant on computational models and simulation results to predict their performance, safety, and reliability. Although important decisions are made based on CSM, the credibility (or trustworthiness) of these models and simulation results is not often questioned by the general public, the technologists who design and build the systems, or the decision makers who commission their manufacture and govern their use.

What is the basis for this trust? The public and decision makers do tend to trust graphical and numerical presentations of results that are plausible and make sense, but their trust is founded on faith in the knowledge and abilities of the engineers and scientists who develop, exercise, and interpret the models. Those responsible for the computational models and simulations on which society depends so heavily are, therefore, keepers of the public trust with an abiding responsibility for ensuring the veracity of their simulation results.

Engineers and scientists are aware that the computational models they develop and use are approximations of reality and that these models are subject to the limitations of available data, physical theory, mathematical representations, and numerical solutions. Indeed, a fundamental approximation in solid mechanics is modeling the nonhomogeneous microstructure of materials as a mathematical homogeneous continuum. Another approximation that is commonly made includes assuming the sections of a beam remain plane during bending. Additionally, a significant approximation that must be made is the characterization of complex material behavior subject to extreme conditions. The use of these approximations, along with their attendant mathematical formulations and numerical solution techniques, has proved to be a convenient and acceptably accurate approach for predicting the behavior of many engineered structures.

Modelers need to ensure that their approximations of reality are appropriate for answering specific questions about engineered systems. The primary goal for the modeler is to establish that the accuracy of the computational model is adequate for the model's intended use. The required accuracy is related to the ability of a simulation to correctly answer a question—ranging from a qualitative question that requires a simple "yes" or "no" response to a quantitative question that requires a numerical value in response. Accuracy requirements vary from problem to problem and can be influenced by public perception and economic considerations as well as by engineering judgment.

The truth of a scientific theory, or of a prediction made from such a theory, cannot be proved using deductive logic. However, scientific theories and subsequent predictions can and should be tested for trustworthiness by the accumulation of evidence. The evidence collected, corroborative or not, should be organized systematically through the processes of computational model V&V. V&V/UQ addresses the issue of trustworthiness by providing a logical framework for accumulating and evaluating evidence and assessing the credibility of simulation results to answer specific questions about engineered systems.

2.1 Purpose and Scope

The purpose of this Standard is to provide the CSM community with a common language, a conceptual framework, and general guidance for implementing the processes of computational model V&V. To this end, this Standard includes a glossary of terms, figures illustrating the recommended overall approach to V&V activities, and discussions of factors that should be considered when developing and executing a V&V program. In creating this Standard, the ASME V&V 10 Subcommittee benefited from the earlier contributions to the field of V&V by other groups, especially AIAA (ref. [2]), as well as by Oberkampf et al. (ref. [3]) and Thacker et al. (ref. [4]).

Three related documents (see Preface) have been or are in the process of being generated that provide details not presented in this overarching Standard. ASME V&V 10.1 describes a simple example of V&V to illustrate some of the key concepts and procedures presented in this Standard; ASME V&V 10.2 will address uncertainty within the V&V process; and ASME V&V 10.3 will provide an in-depth description of some of the metrics that can be used for validation comparisons and how to apply them.

To maximize the value to the engineering community, the ASME V&V 10 Subcommittee chose to write this Standard from the perspective of V&V for high-consequence computational predictions of complex engineering systems. However, the guidance provided here is also appropriate for simple applications, though it is understood that smaller budgets and lower risks will reduce the scope of the V&V effort. While the concepts and terminology presented here are applicable to all applied mechanics, the focus is on CSM.

2.2 General Concepts of V&V

2.2.1 Definitions. Some basic terms that provide the basis for the rest of this Standard include

(a) Code. A code is the computer implementation of algorithms developed to facilitate the formulation and approximate solution of a class of problems.

(b) Model. A model is the representation of a system, phenomena, or process under specific physical conditions. The representation includes conceptual, mathematical, and computational models.

(c) Simulation Results. Simulation results are raw or processed calculations obtained by running the computational model.

(d) Verification and Validation. The terms “verification” and “validation” have been used interchangeably in casual conversation as synonyms for the collection of corroborative evidence. The definitions used in this Standard are largely consistent with those published by the DoD [ref (1)] and the AIAA [ref. (2)].

(1) Verification is the process of determining that a computational model accurately represents the underlying mathematical model and its solution.

(2) Validation is the process of determining the degree to which the model is an accurate representation of corresponding physical experiments from the perspective of the intended uses of the model.

Additional terms that form part of the shared language for V&V as used herein are found in Mandatory Appendix I.

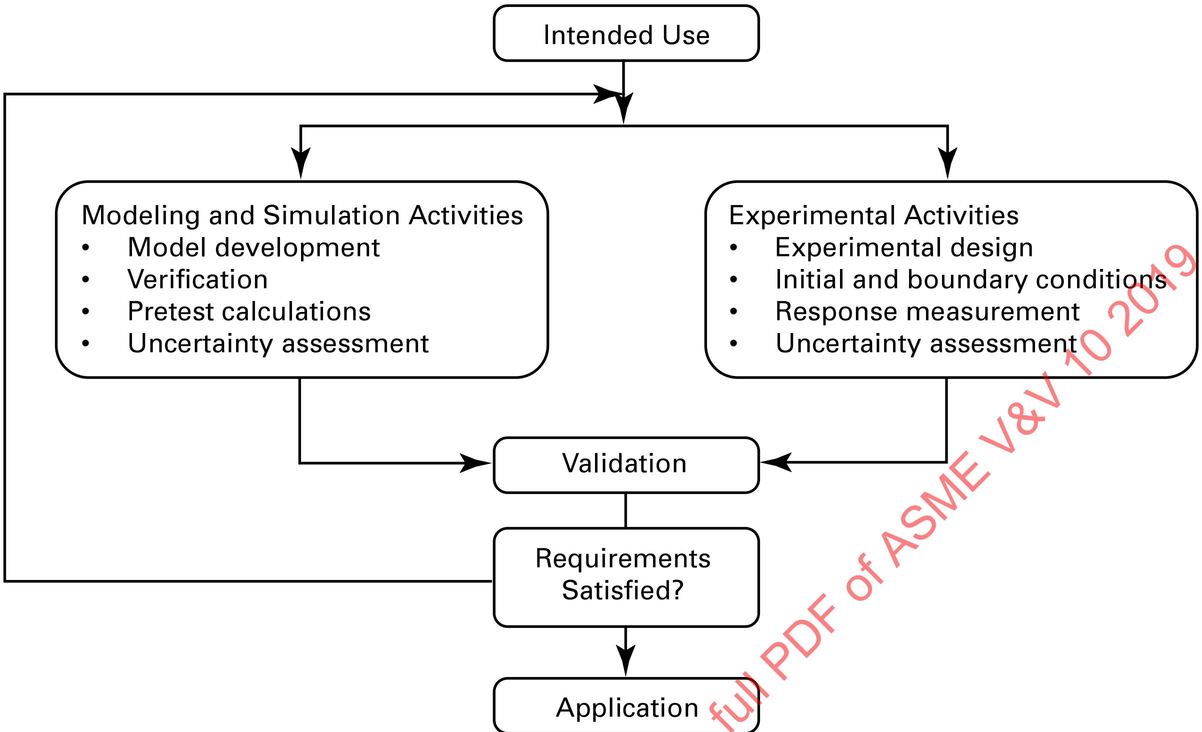
In essence, verification entails gathering evidence to establish that the computational implementation of the mathematical model and its associated solution are correct. Validation, on the other hand, entails comparing simulation outputs with experimental outputs to establish evidence that the appropriate and adequate models were used to answer the questions of interest and to quantify the uncertainties within the process. Validation is attained through meeting the criteria established specifically for determination of validation, i.e., acceptable agreement is obtained.

2.2.2 Objectives. The general objectives of V&V are to assess the reliability of the computer software and numerical methods used in the simulation and assesses the accuracy of the simulation with respect to available experimental observations. The model builder considers the model validated for those response quantities at the experiment locations within the parameter space once predetermined requirements for demonstration of accuracy are met. For the decision maker or other stakeholder, the intended use also defines limitations on the applicability of the model.

An example of an intended use is to predict the response of a particular make and model of automobile in frontal impacts against a wall at speeds up to 30 mph. Validation might consist of numerically simulating the compaction of the front end and the acceleration of the occupant compartment to within 20% for tests at 10, 20, and 30 mph. The model could then be used to predict response quantities at other scenarios across the parameter space. A predicted response for other makes or models of automobiles, for higher speeds, or for rear-end or side collisions would be away from the locations at which the model was validated. For some of those cases, the predicted response would be very far away from the validation points, which are described in greater detail in para. 2.3. Scenarios farther away from the validation location within the parameter space typically have a higher level of uncertainty and a corresponding lower level of confidence in their predicted response. Requirements for accuracy and predictive capability would have to reflect the separation between the validation points and the intended use point.

A detailed specification of the model’s intended use should include a definition of the criteria by which the model’s predictive capability will be assessed. The criteria should be driven by application (i.e., intended use) requirements. For instance, in the previous example, 20% accuracy at the validation points is based on consideration of how the predictions will be used. Although criteria and other model requirements may have to be changed before, during, or after validation

Figure 2.3-1 Elements of V&V/UQ Activities



assessments of the entire system, it is best to specify validation criteria prior to initiating model-development and experimental activities in order to establish a basis for defining how "good" is good enough.

The recommended approach to model V&V emphasizes the need to develop a plan for conducting the V&V program. For complex, high-consequence engineered systems, the initial planning should be done by a team of experts. The V&V plan should be prepared before any validation experiments are performed, because it should guide how the validation tests are conducted and define the specifications of those tests. The plan should include, at a minimum, the following:

(a) a detailed specification of the intended use of the model to guide the V&V effort

(b) a detailed description of the full physical system and the hierarchy into which the system has been decomposed, including the behavior of the system's parts both in isolation and in combination

(c) a list of the experiments to be performed for both calibration and validation

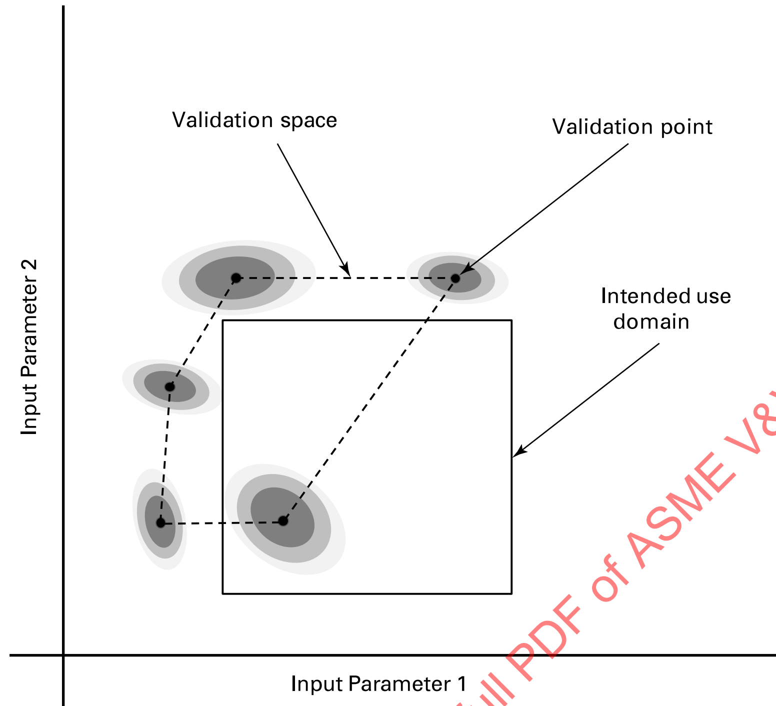
(d) information about how the V&V approach relates to program factors such as schedule, cost, and available resources

Key considerations in developing the V&V plan are discussed in Section 3, following presentation of the V&V approach and process.

2.3 General Concepts of Predictive Capability

Prediction is defined herein as the use of a model to calculate a response where the modeler does not know the experimental outputs. By this definition, a prediction can be made either during the validation process and then compared to experimental results, or after the validation process is complete where no experimental results are, or are expected to be, available. In the validation process shown in Figure 2.3-1, the Modeling and Simulation Activities produce simulation results and the Experimental Activities produce experimental measurements. The process of validation compares the simulation results to the experimental measurements using a validation metric to quantify the difference between the two. Once acceptable agreement is attained (i.e., a validated model has been obtained), that model may then be used to generate simulation results for input conditions that are different from what has been tested. This illustrates the definition of a prediction without experimental results referred to above, and is clearly separated from the simulation performance that occurs during the validation process.

Figure 2.3-2 Relationship Between Validation Points, Validation Space, and Intended Use Domain



A model is validated against experimental results at a specific set of input conditions, which can be referred to as the validation point. The input conditions are generally uncertain; therefore, the validation point will have uncertainty associated with it as well. This concept is illustrated in Figure 2.3-2 for the case of a two–parameter input space, where the validation points are represented by solid black dots and the associated input uncertainties by concentric shaded ellipses. The region encompassed by the validation points is defined as the validation space and is conceptually similar to an interpolation region in regression analysis. The boundary of this space, marked by dashed lines in the figure, illustrates the extent of the parameter space that has been experimentally investigated.

Generally, the parameter space is far more than two dimensions, creating a high-dimensional validation space that can make it difficult to determine where a point of consideration is with respect to the entire validation space. Such high-dimensional spaces are typically nonintuitive and difficult or impossible to visualize. It can also be difficult to determine the correlation structure of a large number of input parameters in high-dimensional spaces. There are other uncertainties not shown in this figure; for example, the uncertainty in the measurement of the system response and the uncertainty in the validation metric.

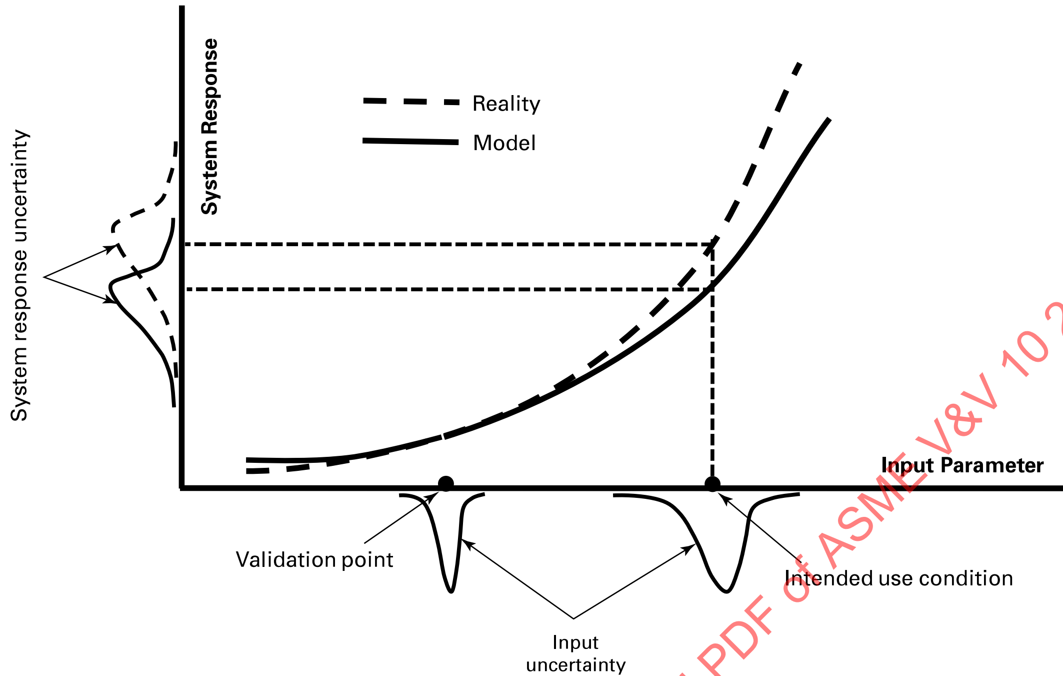
Models are typically developed to make predictions over a range of input conditions. In general, there is an "intended use" domain, and the validation points can fall inside or outside of this domain. From a qualitative standpoint, the accuracy of predictions made with the model depend on the

(a) assumptions and approximations made in the formulation of the mathematical model
(b) number and location of the validation points relative to the location of the intended use domain
(c) degree of uncertainty associated with measurement of the system response at each validation point
(d) input uncertainties that exist in the intended use domain

Engineering experience or intuition may suggest that predictions within the validation space are more reliable than predictions made outside the validation space. How to mathematically quantify this is problematic because the reliability of predictions depends on how quickly the system response changes away from the validation points, yet for many well-behaved problems this suggestion holds true.

How predictive accuracy relates to reality is represented in Figure 2.3-3. The graph shows the prediction ("Model") and the actual system response ("Reality") as functions of a single input parameter. A single validation point is shown as well as a single intended use condition. Input parameter uncertainty is shown for both points; however, the validation condition uncertainty is typically smaller than our knowledge of the intended use condition uncertainty. The highest levels of model accuracy typically occur at the validation points. As one moves away from the validation points, the confidence in

Figure 2.3-3 Depiction of the Increase in Uncertainty for Model Predictions Away From Validation Points



the model accuracy commonly decreases, as would be expected. The divergence between the predicted response and the real response, which is unknown away from validation points, represents the increasing uncertainty due to both systematic bias in the model and random variability in the input data. While a more reliable prediction should be possible within the validation space (as suggested by engineering experience or intuition), it is not guaranteed. A comprehensive determination of the uncertainties of the problem, including both experimental and modeling, are key in determining the predictive accuracy of the model.

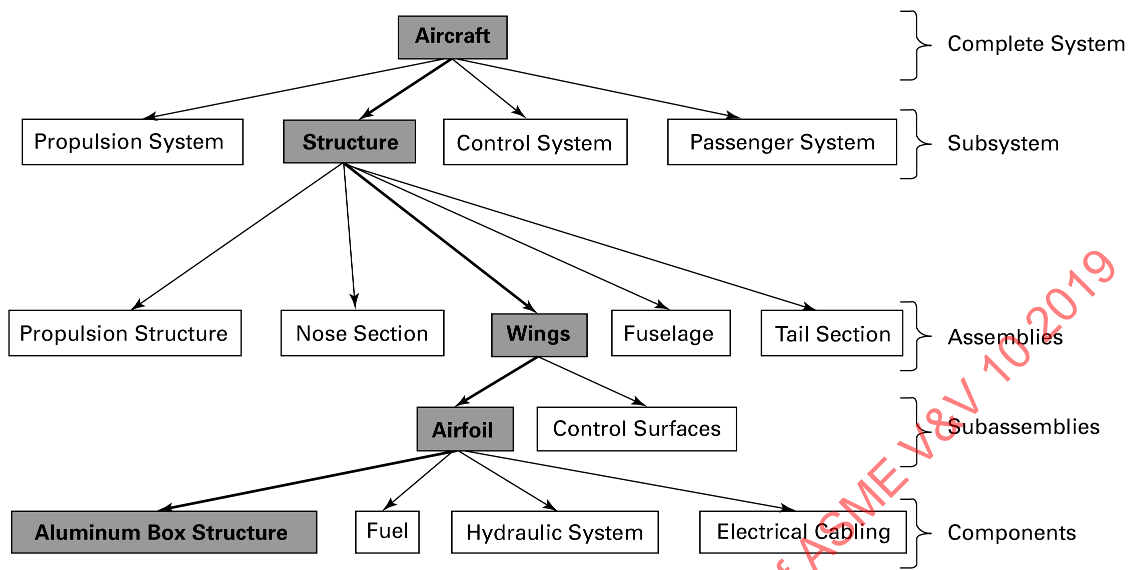
The decision maker using the simulation results must factor in the potential adverse consequences if the predictive capability proves to be unreliable relative to the requirements of system performance, safety, and reliability. These adverse consequences could be associated with corporate liability, loss of potential future business, environmental impact, and public safety. This higher-level decision-making process is beyond the scope of this document and will not be addressed further.

3 APPROACH

3.1 Modeling Complex Systems

Many real-world physical systems that would be the subject of model V&V are inherently complex. To address this complexity and prepare a detailed description of the full system, it is helpful to recognize that the system being modeled is hierarchical in nature. As illustrated in Figure 3.1-1, the hardware of a physical system is typically composed of subsystems, which contain assemblies. Each assembly consists of two or more subassemblies; a subassembly, in turn, consists of individual components. Each separate section of the hierarchy can be seen as a validation case, or an instance that needs to be validated. The top-level validation case in Figure 3.1-1 can be viewed as any level of a real physical system. For example, it could be a complete aircraft, or it could be the wing of an aircraft. If an aircraft is the top-level validation case, it might be composed of subsystems such as the propulsion system, the structure/body, the control system, and the passenger system. Considering the structure as a subsystem, it might be composed of assemblies like the propulsion structure, the nose section, the wings, the fuselage, and the tail section. Similarly, an assembly such as the wings might contain subassemblies like the airfoil and the wing control surface, each of which is composed of components. Each of these subsets at all levels of the hierarchy can be considered a validation case that could be subjected to the validation process. In terms of V&V, the requirements for the model for the top-level validation case, as well as for all lower levels, depend on the intended use of the model.

Figure 3.1-1 Hierarchical Structure of Physical Systems



3.2 Hierarchical Approach to V&V

A top-down decomposition of the physical system into its hardware constituents, as discussed in para. 3.1, is the basis for developing a model of this system. However, the recommended approach to V&V is to develop such a hierarchy and then work from the bottom up, beginning at the lowest tier (i.e., the component level), to identify and describe the physical phenomena at each level that must be accurately simulated with the model. This bottom-up approach recognizes that, while some of the physical responses of components may be representative of a single physical phenomenon (such as deformation, natural frequencies, or buckling loads), at higher levels of the hierarchy, interaction effects not exhibited by the individual components are likely (such as effects of frictional interfaces and joints). For example, a model of a subassembly consisting of a welded automobile frame could introduce behavior that is not present when individual struts are modeled separately.

Building a model from the bottom up will result in a multitiered set of individual models (a system-level model and its embedded submodel[s]) and form the basis for defining validation experiments that need to be conducted at each tier of the hierarchy to ensure that the constituent models at each particular tier function appropriately. Models for components, subassemblies, assemblies, and subsystems that have been validated previously can and should be reused if the response mechanisms they have exhibited and the predictive accuracy they have demonstrated clearly meet the requirements of the new system.

Figure 3.1-1 depicts an overview of the hierarchical approach to validation. The figure identifies the models that could be constructed at each tier, and highlights one potential path through the hierarchy starting with the “Aluminum Box Structure” at the component level and culminating with the “Aircraft” at the complete system level. In this example, validation of the system model will be achieved by consensus of the program experts if the responses of the complete vehicle in laboratory or field experiments are successfully predicted.

The highest-tier validation experiments are typically either special cases of the expected operating conditions or idealized versions of the real world system. It is important to complete V&V with computations and experiments at the system level to assess whether the bottom-up approach adequately considered complex nonlinear interactions at all levels of the hierarchy (i.e., that the appropriate hierarchical decomposition was used). It may be tempting to perform validation of system models directly from data taken from tests of the complete system without new or archived validation at lower levels in the hierarchy. This can be problematic for a large number of components or if the subsystem models contain complex connections or interfaces, energy dissipation mechanisms, or highly nonlinear behavior. If there is poor agreement between the simulation results and the experiment, it is often difficult, if not impossible, to isolate which subsystem model is responsible for the discrepancy. Even if good agreement between calculation and experiment is observed, it is still possible that the model quality is poor because of error cancellation among the subsystem models. A better strategy is to conduct a sequence of experiments that builds confidence in the model’s ability to produce accurate simulations at multiple levels in the hierarchy.

3.3 V&V Activities and Products

Once the elements of the physical system's hierarchy (whether one or many tiers) have been defined and prioritized, a systematic approach can be followed for establishing and increasing confidence in model predictions through the logical combination of hierarchical model building, focused laboratory and field experimentation, and uncertainty quantification. This process is discussed in this subsection.

Figure 3.3-1 illustrates the V&V process, identifying important steps and showing the relationships between the various aspects. This process assumes the hierarchy has been decomposed into individual validation cases and the process will be repeated for each of those cases. The V&V exercise is employed while progressing through the hierarchy toward simulation of the top-level model and its validation experiment(s). This succession of activities will generate evidence to assess confidence in subsequent predictions made using the top-level model for its intended use.

Activities are denoted by simple text, such as "Uncertainty Quantification"; the products of these activities are highlighted in rounded boxes (e.g., the "Simulation Results" are the product of the "Calculation" activity). The outlines for the modeling and simulation and the physical experimentation branches parallel each other throughout the process. Modelers follow the left branch to develop, exercise, and evaluate the model. Experimenters follow the right branch to obtain the relevant experimental results via physical testing. Modelers and experimenters collaborate throughout the process in developing the conceptual model, conducting preliminary calculations for the design of experiments, specifying initial and boundary conditions for calculations for validation, and developing the validation experiments.

The process shown in Figure 3.3-1 is repeated for each element of every tier of the hierarchy (i.e., every validation case) in the system decomposition exercise discussed in para. 3.2, starting at the component level and progressing upward to the complete system. In a bottom-up approach, both preliminary conceptual model development and V&V planning for all levels in the hierarchy, especially the system level, are performed before the main validation activities for components, subassemblies, assemblies, and subsystems begin to establish any interdependencies that may exist. Results of each completed validation case are incorporated into the V&V of the top-level system, and then the next validation case is addressed. This loop is repeated until the complete system has been exercised through the process. At that point the validation process is complete and simulations for conditions of intended use of the system model can be performed.

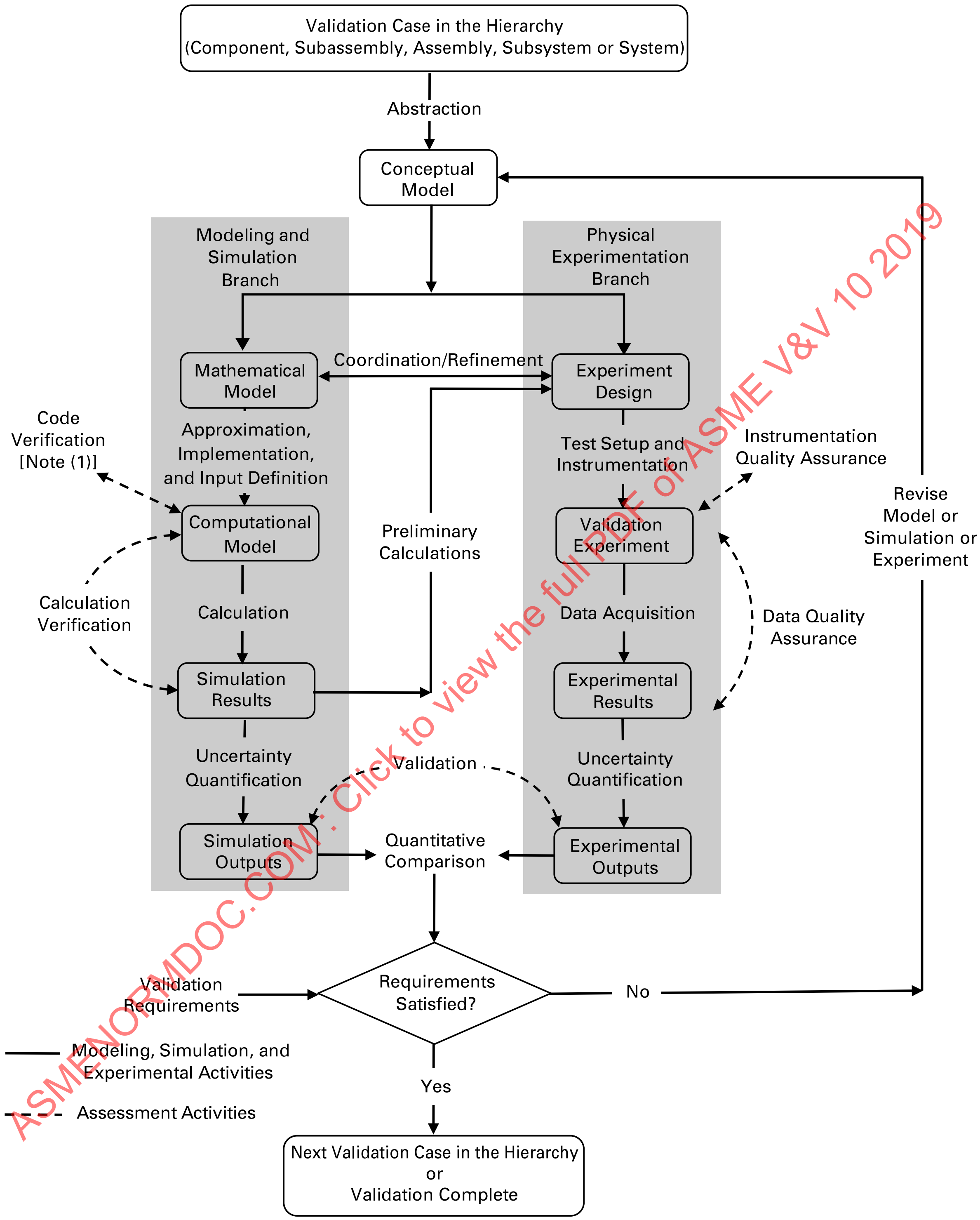
Abstraction of each validation case into the conceptual model requires identifying the domain of interest, important physical processes and assumptions, and response quantities of interest. The abstraction essentially produces the modeling approach based on these considerations. It is also intimately connected to the development of the overall V&V plan that establishes the validation requirements, including the types of experiments to be performed and the required level of agreement between the experimental outputs and the simulation outputs. Thus, this activity is typically iterative and involves a combined effort by modelers, experimenters, and decision makers.

3.3.1 The Modeling and Simulation Branch. Through idealization, the modeler constructs a mathematical interpretation of the conceptual model. The resulting mathematical model is a set of equations and modeling data that describe physical reality, including the geometric description, governing equations, initial and boundary conditions, constitutive equations, and external forces. During the subsequent approximation, implementation, and input definition activity, the modeler develops the computational model, which is the software implementation on a specific computing platform of the equations developed in the mathematical model, usually in the form of numerical discretization, solution algorithms, and convergence criteria. The computational model includes numerical procedures, such as finite element or finite difference, for solving the equation prescribed in the mathematical model with specific computer software.

Modelers do not always develop and implement equations to create new computational models; instead, they often use existing ones (e.g., commercial codes, graphical user interfaces, etc.). It is important for modelers to consider the steps described herein when using these existing computational models and carefully think about any potential challenges arising from them. The modeler must ensure that tools he or she has not personally developed are specifying the model he or she really intends.

In the code verification assessment activity, the modeler uses the computational model to assess a separate set of problems with known solutions. These problems typically have much simpler geometry, loads, and boundary conditions than the validation problems to identify and eliminate algorithmic and programming errors. This assessment activity is applied not to the validation case, but rather to established problems with known solutions. In the subsequent calculation verification activity, the version of the computational model used for the validation case (i.e., with the geometries, loads, and boundary conditions typical of that problem) is used to identify sufficient mesh resolution to produce an adequate solution, including the effects of finite precision arithmetic. Calculation verification yields a quantitative estimate of the numerical precision and discretization accuracy for calculations made with the computational model for the validation experiments. In the calculation activity, the modeler runs the computational model to generate the simulation results for the validation case. The simulation results can also be post-processed to generate response quantities for comparison with experimental results. A response quantity can be as simple as the maximum response for all times at a specific

Figure 3.3-1 V&V Process



NOTE: (1) Code verification is performed using different models with closed-form or manufactured solutions.

location in the object being tested, or as complex as a fast Fourier transform of the complete response history at that location.

In the subsequent UQ activity, the modeler should quantify the uncertainties in the simulation results that are due to inherent variability in model parameters or lack of knowledge of the parameters or model form. The results of the parameter and model-form UQ should be combined with those of the calculation verification to yield an overall uncertainty estimate associated with simulation results. Response quantities of interest extracted from simulation results and estimates of uncertainty combine to form the simulation outputs that are used for comparison with the experimental outputs.

3.3.2 The Experimental Branch. The experimenter first idealizes a physical representation of the conceptual model and how to construct it. The resulting experimental design is a set of material specifications, boundary conditions, initial conditions, and instrumentation requirements that are necessary to observe and measure the effect that changes to input variables have on the solid mechanics behavior of the validation case. The purpose of validation experiments is to provide information needed to assess the accuracy of the model; therefore, all assumptions should be understood, well defined, and controlled.

During the planning of the experiment, preliminary calculations (including sensitivity and uncertainty analyses) are recommended to assist with the design of the experiment by, for example, identifying the most effective locations and types of measurements required. These data should include not only response measurements, but also measurements needed to define model inputs and model input uncertainties associated with loading, initial conditions, boundary conditions, etc.

The modeler and the experimenter should work together so that each is continually aware of assumptions in the models or the experiments. By observing the preparations for the experiment, for example, the modeler may be able to detect incorrect assumptions in the model. However, experimental results should not be given to the modeler to preclude inadvertent or intentional tuning of the model to match experimental results.

Implementing the experiment design into a validation experiment involves setting up the physical articles, installing instrumentation, and confirming the setup. Instrumentation calibration is the process of evaluating the accuracy of the instrument before and after the experiment. It is necessary to perform instrumentation quality assurance to ensure the correctness of the data collection process; this is similar to the code verification performed on the modeling side of Figure 3.3-1.

Instrumentation quality assurance is an assessment of whether the measurement system is acceptably accurate and repeatable for its intended purpose. This is often called gage repeatability and reproducibility (gage R&R). When conducted properly, these studies identify the measurement variation attributable to various parts of the complete measurement system, such as the instruments, the operator, or the physical process under study. Performing these studies helps identify the anticipated levels of variation and, if those prove unacceptably high, target the areas where improvements can be made. With adequate measurement systems, most of the variation would be expected within the test article itself and not the operator or the instrumentation. It is beyond the scope of this document to lay out specific instrument quality assurance and gage R&R approaches, but there is a wealth of resources that address this topic (ref. [5]).

Data acquisition involves the collection of raw data from various instruments used in the experiment (e.g., strain and pressure gages and high-speed cameras) and the generation of processed data (e.g., time integrals, averages, or the determination of velocity from high-speed video). The experimental results can be transformed as necessary into experimental quantities that are more useful for direct comparison with simulation outputs. Multiple experiments are generally required to quantify uncertainty due to inherent variability.

Data quality assurance (Data QA) is an assessment of whether data captured during a validation experiment are a reasonable representation of what occurred during the experiment. Many different factors can cause discrepancies at this stage of the process, including but not limited to loss of network connections, operator error, instrument saturation, misalignment, and interference. Data QA represents a quality check on the data collected, not an in-depth analysis of results. Many methods exist for providing basic checks on the quality and consistency of data and some of the most effective are graphing or visualizing the data and searching for missing, out-of-range, or other impossible data, such as error code in a numeric field or negative values in a timestamp. It is also important to search for outliers and other extreme or improbable values not necessarily to eliminate them but to identify them as close to the time of capture as possible to preserve context for further investigation.

Basic statistical summaries should be performed on the data to determine if the expected levels of variability exist. The presence of little to no variability in experimental measurements can be an indicator of poor connectivity to or functioning of experimental instruments. Statistical summaries before and after processing should be compared for consistency if data aggregation from multiple devices, transformations, or other processing occurred.

The experimenter should next perform UQ to quantify the effects of various sources of uncertainty on the experimental results. These sources include measurement error, design tolerances, manufacturing and assembly variations, unit-to-unit fabrication differences, and variations in performance characteristics of experimental apparatuses. Experimental outputs, which are the product of this UQ activity, will typically take the form of experimental results plus quantified uncertainties as a function of time or load.

3.3.3 Assessing Agreement. Once experimental and simulation outputs for the actual test conditions have been generated, the modeler and experimenter perform the validation assessment activity by comparing these two sets of outputs.

The metrics for comparing experimental and simulation outputs as well as the criteria for meeting the requirements will have been specified during the formulation of the V&V plan. The essential result of model validation assessment is the quantitative assessment of the model's ability to predict the experimental results obtained.

The diamond symbol asking "Requirements Satisfied?" near the bottom of Figure 3.3-1 provides an objective decision point for initiating improvements in the conceptual, mathematical, and computational models and in the experimental designs.

The block at the bottom of Figure 3.3-1 denotes that the process repeats for the next submodel to be validated, either at the same tier or at the next higher tier of the hierarchy. Thus, as V&V is performed, the results of the component-level activities (including the uncertainties) are aggregated and propagated to the next higher tier of the hierarchy, and so on up to the full-system level. Once all of the validation cases have been addressed (i.e., validation performed at all levels of the hierarchy that were specified in the validation plan), then the validation process for that specified hierarchy is complete.

3.4 Development of the V&V Plan

A V&V program should be thoughtfully planned before the major activities in model development and experimentation are initiated. In particular, it is essential to define the requirements for system-level validation in the V&V plan.

3.4.1 Validation Testing. In many instances, the most difficult part of V&V planning is establishing the relationship between the validation experiments and the intended use cases. It may not be possible either to test the complete system or to test subsystems and assemblies over the full range of conditions of interest. For example, when modeling the response of a complete aircraft, it is unlikely that all of the important response quantities will be available for all of the flight conditions of interest. Still, a plan that defines the set of conditions for which the systems, subsystems, and assemblies should be tested at validation conditions of interest should be developed by a consensus of experts. These conditions must be balanced by what is achievable with the given resources and time constraints.

3.4.2 Selection of Response Quantities. Complex physical systems and their corresponding model simulations can encompass an enormous array of response quantities. Because only a limited number of measurements can be made in validation experiments, it is important to identify the response quantities of interest before the experiments are designed or the models developed. The selection of which response quantities to measure and compare with predictions should be driven by application requirements. At the system level, this may require product safety or reliability parameters to be defined in engineering terms. For example, occupant injury in automobile crashes may be related to occupant–compartment accelerations and protrusions, and thus those quantities should be measured and predicted. The appropriate response quantities of the other levels of the system hierarchy depend on how their responses affect the critical quantities of the system response. Specifications should also be made for the metrics used for comparisons of outputs, such as root-mean-square differences of simulation and experimental acceleration histories.

3.4.3 Accuracy Requirements. The accuracy requirements for predicting the response quantities of interest with the system-level model are based on the intended use and may rely on engineering judgment or a formal risk analysis. Specification of accuracy requirements allows the question of acceptable agreement to be answered quantitatively. Only with accuracy requirements can the decision be made to accept or revise a model. Without accuracy requirements, the question of "How good is good enough?" cannot be answered.

System-level accuracy requirements are used to establish accuracy requirements for each submodel in the V&V hierarchy. These requirements should be established such that models for subsystems, assemblies, subassemblies, and components are refined at least to the degree required to meet the accuracy goal of the system-level model. A sensitivity analysis of the complete system can be used to estimate the contribution of each model; the estimated contributions can then be used to establish commensurate accuracy requirements. It is reasonable to expect that the accuracy requirements for component behavior will be more stringent than the accuracy requirements for the complete system, due to the simpler nature of problems at the component level and the compounding effect of propagating inaccuracy up through the hierarchy. For example, a 10% accuracy requirement might be established for a model that calculates the axial buckling strength of a tubular steel strut in order to achieve 20% accuracy of the collapse strength of a frame made of many such components.

3.5 Documentation of V&V

It is important to document both the rationale and the results of V&V not only for the current intended use but also for future potential uses. V&V allow a knowledge base to be built from the various levels in the hierarchy and then reused in subsequent applications. For example, in many applications, derivative or closely related product designs are used in the development of future designs. If a thorough execution and documentation of hierarchical V&V has been performed for the model of the basic design, many of the hierarchical elements for V&V of the model for the derivative design might be reusable. In this way, the value of investment in hierarchical V&V can be leveraged to reduce V&V costs for future projects. Documentation also provides the basis for possible limitations on reuse and thus prevents unjustifiable extrapolations. The V&V documentation should be comprehensive, self-contained, retrievable, and citable.

3.6 Overview of Subsequent Sections

Sections 2 and 3 have outlined the basic principles and characteristics of a careful and logical approach to implementing model V&V for CSM. The guidelines for accomplishing the various activities in V&V form the contents of sections 4 through 6. Model development activities are the focus of section 4. In section 5, the two assessment activities of code verification and calculation verification are described. Section 6 discusses the experimental and assessment activities involved in validating a model. The concluding remarks in section 7 identify issues that need to be addressed so that V&V for CSM can evolve into a more robust and quantitative methodology. The concluding remarks are followed by a glossary of V&V terms (Mandatory Appendix I) and references (Mandatory Appendix II).

4 MODEL DEVELOPMENT

This section describes the activities involved in computational model development, starting with formulating the conceptual and mathematical models and then revising these models during V&V and, finally, quantifying the uncertainty in the resulting model. Model development activities begin with the assumption that the validation case, the intended use of the model, the response quantities of interest, and the accuracy requirements have been clearly defined for that particular model. There will be some interplay between the development of the conceptual model and the V&V plan. In general, the system model (conceptual to computational) is built up from subsystem, assembly, subassembly, and component models, as illustrated in Figure 3.1-1. At the highest level of the hierarchy, the "validation case" within Figure 3.3-1 is the real-world system, assuming that the experiment is conducted with the goal of model validation. However, as discussed in para. 3.4.1, this is commonly not possible. In that situation, it is necessary to explicitly include the estimated uncertainty in the prediction and rely on the predictive capability of the simulation.

Figure 4-1 illustrates the path from a conceptual model to a computational model. An example of a conceptual model is a classic Bernoulli-Euler beam with the assumptions of elastic response and plane sections. This conceptual model can be described with differential calculus and other mathematical assertions to produce a mathematical model. The equations can be solved by various numerical algorithms, but in CSM they are typically solved using the finite element method. The numerical algorithm is programmed into a software package, here called a "code." With the specification of physical and discretization parameters, the computational model is created.

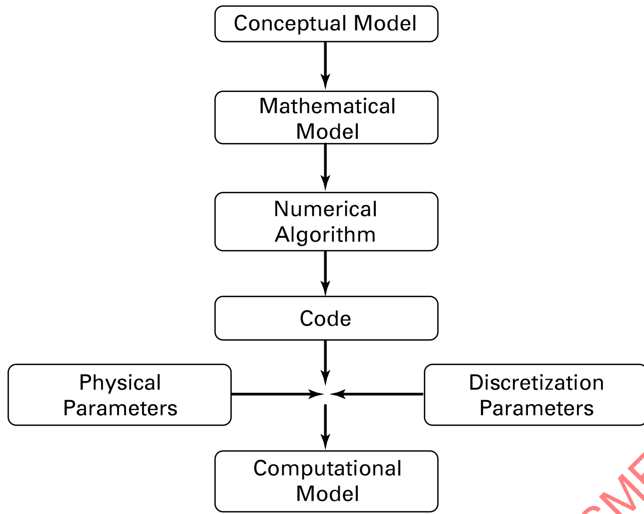
4.1 Conceptual Model

The conceptual model is defined as the idealized representation of the solid mechanics behavior of the validation case. This model should therefore include those mechanisms that affect the key mechanical and physical processes that will be of interest for the intended use of the model. Conceptual model development involves the formulation of a mechanics-based representation of the validation case that is amenable to mathematical and computational modeling, includes the appropriate level of detail, and is expected to produce results with adequate accuracy for the intended use. Essentially, it defines the modeling approach.

The formulation of the conceptual model is important to the overall model-development process because many fundamental assumptions that influence interpretation of the simulation results are made at this stage. These assumptions include the

(a) determination of how many separate parts or components will be included in the model
(b) approach to modeling the material behavior
(c) elimination of unimportant detail features in the geometry
(d) selection of interface and boundary types (e.g., fixed, pinned, contact, friction, etc.)

Figure 4-1 Path From Conceptual Model to Computational Model



If an important mechanical phenomenon is omitted from the conceptual model, the resulting simulations might not be adequate for the intended use of the model.

An essential step in developing the conceptual model is to identify which physical processes within the validation case are anticipated initially to have significant effects on the system's response. Likewise, it is important to identify which physical processes do not have a significant effect and to note that such mechanics will be ignored in the conceptual model. Identifying the essential physical processes will help to ensure that the computational model sufficiently represents the mechanics involved and does not waste computational effort modeling physical effects that do not affect the response quantities of interest. Development of the conceptual model also requires knowledge of the range of operating environments that are relevant to the model's intended use. The environments affect choices in the modeling, such as whether to include plasticity or thermal softening.

Response quantities are the characteristics of the response of the physical system that the computational model has to predict for the intended use. They could include characteristics such as the maximum tensile stress in bolts, the peak acceleration of the center of a floor, the average value of pressure in a chamber, the deflection of the center of a glass window, the modal frequencies of a radio tower, or the strain energy release rate at the tip of a fracture. Knowledge of the response quantities is important in the conceptual modeling activity because interest in certain response quantities may influence decisions that are made during the mathematical and computational modeling activities. For example, if the deflections of a particular part are of interest, the compliance of materials surrounding that part should not be neglected.

During development of the conceptual model, the best tools available for identification of the key physical processes are engineering expertise and judgment. Thorough documentation of the rationale for what is included in — or excluded from — the conceptual model is an important part of proper model development and traceability. Note that once the computational model has been developed, a sensitivity analysis can be used to investigate the importance of a physical process to the response of the system (see para. 4.5).

Constructing a Phenomena Identification and Ranking Table (PIRT) is useful for identifying the key physical processes (ref. [6]). The PIRT is both a process and a product. The process involves gathering a diverse group of subject-matter experts together to rank the physical phenomena according to their importance to the system response quantities of interest. The product is the table itself, which presents a summarized list of the physical phenomena along with a ranking (e.g., high, medium, low) of the importance of each phenomenon to the system response quantities of interest. Sample entries in a PIRT are shown in Table 4.1-1. The PIRT can be used either to construct a conceptual model (starting from scratch) or to prioritize the conceptual model of a large general-purpose code that may have the ability to model hundreds of phenomena, only a subset of which are relevant to the subject model.

At this stage of model development, the PIRT can include a qualitative judgment regarding the ability of either existing or to-be-developed computational models to describe the physical processes accurately (see the last column in Table 4.1-1). This information helps prioritize which physical processes will be investigated experimentally during validation (i.e., it is part of the interplay between the development of the conceptual model and the development of the V&V plan). For the example in Table 4.1-1, phenomenon B has a low priority for validation because it can

Table 4.1-1 Phenomena Identification and Ranking Table (PIRT) Example

Phenomenon	Type of Phenomenon	Importance to Response of Interest	Level of Confidence in Model
A	Interface	High	Medium
B	Plasticity	Medium	High
C	Loads	Medium	Low
D	Fracture	Low	Low

already be modeled with high confidence. Similarly, phenomenon D has a low priority because of its low importance to the model response of interest.

4.2 Mathematical Model

The development of the mathematical model consists of specifying the mathematical descriptions of the mechanics represented in the conceptual model. In the mathematical model, principles of mechanics, the material behavior, interface properties, loads, and boundary conditions are cast into equations and mathematical statements. For example, if the property of an interface between two bodies is to be described with Coulomb friction, the mathematical model would be $\tau = \mu\sigma$, where

μ = the Coulomb friction coefficient
σ = the normal stress
τ = the shear stress

The specification of the mathematical model allows the model input parameters to be defined. The model input parameters describe the various user-specified inputs to the model, such as material constants, applied loads, and the Coulomb friction coefficient in the previous example. The domain of interest can then be expressed in terms of these parameters. For example, if the application domain specifies a range of applied loads, a specific parameter (or set of parameters) in the mathematical model can be used to define that range of loads.

4.3 Computational Model

The computational model is the numerical implementation of the mathematical model that will be solved on a computer to yield the computational predictions (simulation results) of the system response. As defined herein, the computational model includes the type and degree of spatial discretization of the geometry (e.g., into finite elements), the temporal discretization of the governing equations, the solution algorithms to be used to solve the governing equations, and the iterative convergence criteria for the numerical solutions. With this inclusive definition, models employing solution-adaptive mesh-generation methods are defined by their adaptive control parameters.

The computational model can be simple or complicated, and it can employ in-house or commercial finite-element software to develop and solve the numerical equations. The modeler may be tempted to jump directly from a geometric description of the validation case to the development of a computational mesh, especially given the availability of highly automated preprocessing software. Meshing, however, is not modeling. The modeler must understand the underlying conceptual model and mathematical model in order to understand the effects on the model outputs that are caused by the assumptions and mathematical simplifications inherent in the computational model. Without this understanding, it is difficult to know whether the computational model is inadequate or inappropriate for the intended use. For example, the modeler must consider the type of boundary conditions to be imposed in buckling problems, because buckling results are sensitive to the end conditions used in the model.

4.4 Model Revisions

At some stage of modeling and simulation, the modeler may find that the computational model needs revisions to achieve the desired accuracy or to account for new requirements. In a general sense, there are two classes of possible revisions to the mathematical and computational models. The first class of revisions covers updates to parameters in the mathematical or computational model that are determined by calibrating the computational model to experimental results (e.g., apparent material parameters, modal damping coefficients for linear vibration, or friction coefficients for a mechanical interface). The second class of revisions covers changes to the form of the mathematical or conceptual model to improve the description of the mechanics of interest so that better agreement with the reference experimental results can be achieved. The two classes of revisions are discussed in paras. 4.4.1 and 4.4.2.

4.4.1 Updates to Model Parameters by Calibration. Revision by parametric model calibration is extensively used in the field of linear structural dynamics to bring computational predictions into better agreement with measured response quantities such as modal frequencies and mode shapes. This revision process is commonly known as "model updating," "model tuning," "parameter calibration," or "parameter estimation." The process allows the most common sources of modeling (and experimental) difficulties in linear structural dynamics — compliance in joints, energy loss/damping, unmeasured excitations, uncertain boundary conditions — to be represented as simple mechanical models and calibrated so that the global response of the computational model is in agreement with the experimental results. Calibration of the model should be performed only after both code verification and calculation verification have been performed.

Parametric model calibration determines only the model's fitting ability, not its predictive capability. A model calibrated to experimental results may not yield accurate predictions over the range of its intended use. This means that the model should not be used as a calibration framework for some uncertain parameters if these parameters can be evaluated in independent tests. *Data used for model calibration must remain independent of data used to assess model validation*.

The type of experiment used to determine the values of unknown or uncertain model input parameters is generally referred to as a "calibration experiment." The goal of a calibration experiment is distinct from the goal of a validation experiment. The purpose of a calibration experiment is to generate values or quantified probability distributions for model input parameters under specific types of experimental conditions. For example, an optimization approach may be used to determine the parameter values using a computational model of the calibration experiment and the measured data from the calibration experiment. In contrast to calibration experiments, validation experiments are designed and performed to provide an independent, objective assessment of the predictive capabilities of the computational model.

It is a reality of modeling, given cost and schedule constraints, that model calibration is often performed after an initial validation assessment has been made and the requirements have not been satisfied (as indicated in Figure 3.3-1). That is, the modeler finds a set of parameter values that provides acceptable agreement with the validation test data, but only after failing to achieve that agreement with a prediction. Unfortunately, to then assess predictive capability (outside of the domain of the validation referent data), subsequent validation against other independent experiments may still be necessary. Any revisions to the parameter values after V&V are applied signifies new model-development activity, triggering a repetition of some model V&V.

4.4.2 Updates to Model Form. The second class of model revisions consists of changes to the form of the conceptual model and, in turn, the mathematical model and the computational model. Typically, the need to revise the model form is observed during the quantitative comparison activity, when some characteristics in the response of the structure are not consistent with the corresponding characteristics of the model output, and the differences are not attributable to reasonable uncertainties in the model parameters.

The following are among the many common types of deficiencies in model form that can be responsible for inaccurate simulation results:

(a) two-dimensional models that cannot represent three-dimensional response effects
(b) inappropriate form for representation of material behavior
(c) assumptions about contacting surfaces being tied when in reality a gap develops between the parts
(d) assumptions that two parts do not move relative to one another when in reality they do, resulting in development of significant friction forces
(e) assumed rigid boundary conditions that turn out to have significant compliance

It is important to look for possible violation of the assumptions of the form of the mathematical model when reconciling the measured data with the results of the computational simulation. As with parameter calibration, any revisions to the model after V&V are applied signifies new model-development activity, triggering a repetition of some model V&V.

4.5 Sensitivity Analysis

Another way, besides intuition and experience, to identify important phenomena is to perform a sensitivity analysis using the computational model. Sensitivity analysis is the general process of discovering the effects of model input parameters on the response quantities of interest using techniques such as analysis of variance (ref. [7]). When performed before the computational model is validated (but not before it is verified), a sensitivity analysis can provide important insight into the characteristics of that computational model and can assist in the design of experiments as part of the PIRT process. Model sensitivities, however, must eventually be subject to the same scrutiny of V&V as the main parameters of interest. As with engineering judgment or even the initial PIRT prioritization, unvalidated model sensitivities may be wrong in magnitude or even in sign (i.e., "+" and "−," or direction). Thus, model sensitivity analysis should be revisited after model revision.

Local sensitivity analysis is used to determine the character of the response quantities with respect to the input parameters in a local region of the parameter space (i.e., in the vicinity of a single point). Finite difference techniques or adjoint methods are used to determine the local gradients at points in the design space. Global sensitivity analysis is concerned with some type of average behavior of the response quantities over a large domain of the parameters and is often used to select a subset of the parameters for detailed local sensitivity analysis.

4.6 Uncertainty Quantification for Simulations

Validation for computational mechanics models must take into account the uncertainties associated with both simulation results and experimental results. The uncertainties associated with experimental results are discussed in section 6. Throughout the modeling process (see the left branch of Figure 3.3-1), and especially during the UQ activity, all significant sources of uncertainty in model simulations must be identified and treated to quantify their effects on predictions made with the model. It is useful to categorize uncertainties as either irreducible or reducible.

4.6.1 Irreducible Uncertainty. Also called "aleatory uncertainty," irreducible uncertainty refers to inherent variations in the physical system being modeled. This type of uncertainty always exists and is an intrinsic property of the system. Examples of irreducible uncertainty are variations in geometry, material properties, loading environment, and assembly procedures. The inherent variability in model parameters is typically characterized by performing replicate component-level tests that cover the range of conditions over which the individual parameters will be exercised in the intended use of the model. If no component-level validation testing is performed, estimates of the inherent variability in model parameters should be based on prior experience and engineering judgment. However, even the most complete set of test information will not eliminate irreducible uncertainty, which can only be better quantified by means such as determining a parameter's mean value, distribution, and distribution form (e.g., normal, uniform, log-normal).

Using probabilistic analysis, inherent variability can be propagated through the simulation to establish an expected variability in the simulation output quantities. Sampling-based propagation methods such as Monte Carlo and Latin Hypercube are straightforward techniques for propagating variability (ref. [8]). Sampling-based methods draw samples from the input parameter populations, evaluate the deterministic model using these samples, and then build a distribution of the appropriate response quantities. Well-known sensitivity-based methods include the first-order reliability method (ref. [9]), advanced mean value (ref. [10]), and adaptive importance sampling (ref. [11]).

4.6.2 Reducible Uncertainty. Also called "epistemic uncertainty," reducible uncertainty refers to deficiencies that result from a lack of complete information or knowledge. Two important sources of reducible uncertainty are statistical uncertainty and model form uncertainty. Statistical uncertainty arises from the use of limited samples. For example, if the mean value of a material property is calculated with only two or three measurements of the material property, then the mean value will contain statistical uncertainty, which can be reduced by considering additional measurements of the material property. Model form uncertainty refers to the uncertainty associated with modeling assumptions and approximations, such as a constant parameter assumption (regardless of its assigned numerical value) in the partial differential equations (PDEs). In other words, a parameter in an equation in the computational model could be defined as having a constant value, whereas in reality the value of the parameter varies with time, temperature, or position. In general, model form uncertainty is extremely difficult to quantify, but some innovative approaches to this problem have been developed (refs. [12], [13]).

4.7 Documentation of Model Development Activities

It is important to document model development activities to facilitate reuse of the model. The documentation should explain the rationale for model development (e.g., modeling assumptions) and describe the conceptual, mathematical, and computational models. The description of the mathematical model should include assumptions about the mechanics of interest and the sources of information for the model parameters. The description of the computational model should include discretization assumptions, computational parameters, and other parameters of interest.

5 VERIFICATION

The process of verification assesses the fidelity of the computational model to the mathematical model. The mathematical model is commonly a set of PDEs and the associated boundary conditions, initial conditions, and constitutive equations. The computational model is the numerical implementation of the mathematical model, usually in the form of numerical discretization, solution algorithms, and convergence criteria. Verification assessments consider issues related to numerical analysis, software quality engineering (SQE), programming errors in the computer code, and numerical error estimation. Verification should precede validation activities because verification deals with the numerical mapping of the mathematical model into a reliable solution that is usable by engineers and scientists.

Verification is composed of two fundamental activities: code verification and calculation verification. Code verification ensures, to the degree necessary, that there are no programming errors in a computer code and that the numerical algorithms for solving the discrete equations yield accurate solutions with respect to the true solutions of the PDEs. Calculation verification estimates the numerical solution errors present in every simulation result; examples include temporal and spatial discretization error, iterative error, and round-off error. Calculation verification is also referred to as numerical error estimation. References [14] and [15] discuss the differences between and emphases of code verification and calculation verification.

Mathematically rigorous verification of a code requires proof that the algorithms implemented in the code correctly approximate the underlying PDEs and the stated initial conditions and boundary conditions. In addition, it would also have to be proved that the algorithms converge to the correct solutions of these equations in all circumstances under which the code is applied. Such proofs are not currently available for general purpose computational physics software. Executing the elements of code verification and calculation verification identified as necessary in this document is critical for V&V, but not sufficient in the sense of mathematical proof (ref. [16]).

5.1 Code Verification

The assessment activity of code verification can be logically segregated into the following two parts:

(a) numerical code verification, which focuses on the underlying mathematical correctness and specific implementations of discrete algorithms for solving PDEs

(b) SQE or software quality assurance (SQA), which addresses such matters as configuration management, version control, code architecture, documentation, and regression testing (ref. [15])

Although CSM code users are typically not directly involved in developing and producing the modeling software they use, it is important that these users provide feedback to the developers to ensure that the best software engineering practices are consistently employed for the codes they use. Otherwise, unnecessary faults in the code may affect simulation results intermittently and unpredictably.

5.1.1 Numerical Code Verification. The objective of numerical code verification is to verify that the numerical solution algorithms are correctly implemented (programmed) in the code and that these algorithms are functioning as intended. Numerical code verification relies on careful investigations of topics such as spatial and temporal convergence rates, iterative convergence rates, independence of numerical solutions to coordinate transformations, and appropriate preservation of symmetry related to various types of initial and boundary conditions. In CSM, the primary solution algorithms are the finite-element method and the finite-difference method. Although the formal (theoretical) order of accuracy of these algorithms may be known from power series expansions of the discrete equations, the observed order of accuracy can be different. Thus, an important part of code verification is determining the observed order of accuracy of the solution algorithm, which is the rate at which the solution asymptotically approaches the exact solution as the discretization is refined. This can be done by comparing two or more computational results with different discretizations to an exact solution and observing the rate of convergence.

Many factors can degrade the observed order of accuracy relative to the formal order of accuracy that is reported as a mathematical feature of an algorithm. These factors include programming errors, insufficient mesh resolution to achieve the asymptotic range, mixed accuracy issues, singularities, discontinuities, contact surfaces, mesh clustering, inadequate iterative convergence, and over-specified boundary conditions (refs. [14], [17]). In verification, all of these reasons for degradation in the order of accuracy are evidence of possible algorithmic or code errors and must be understood.

The primary tasks in numerical code verification are defining appropriate test problems for evaluating the accuracy of the numerical algorithms and assessing the performance of these algorithms on the test problems. Numerical code verification depends on comparing computational solutions to the “correct answer,” which is provided by analytical solutions or highly accurate numerical solutions for a set of well-chosen test problems. The correct answer to a physically meaningful problem can only be known in a relatively small number of simple cases that generally exercise only a limited portion of the code. Fortunately, the method of manufactured solutions (MMS) offers a technique for deriving a mathematically exact solution to a closely related problem in order to exercise all aspects of the code that would be activated by the physical problems of interest.

Because such cases assume a very important role in verification, they should be carefully formulated to provide a comprehensive set of test problems for verification of the code.

Two points must be made regarding the paucity of good benchmarks for complex mathematical models. The first is that some solutions are better than others; therefore, a hierarchy of confidence should be recognized. The following organization of confidence (from highest to lowest) for the testing of algorithms is similar to the one suggested in the AIAA Guide (ref. [2]) and is advocated:

(a) exact analytical solutions (including manufactured solutions)

(b) semianalytical solutions (reduction to numerical integration of ordinary differential equations [ODEs], etc.)
(c) highly accurate numerical solutions to PDEs

The second point is that some test problems are more appropriate than others, so application-relevant test problems should be used. These test problems could be ones with which users have a great deal of experience, or they could be ones that are constructed to address specific needs that arise when planning the verification activities.

Paragraphs 5.1.1.1 through 5.1.1.4 provide additional information on the kinds of tests and techniques employed in numerical code verification.

5.1.1.1 Analytical Solutions. Two categories of analytical solutions are of interest in code verification. First, there are those that correspond to plausible — if often greatly simplified — real-world physics. Second, there are manufactured solutions, which are defined and discussed in para. 5.1.1.2. "Physically plausible" analytical solutions are solutions to the mathematical model's PDEs, with initial conditions and boundary conditions that can realistically be imposed, such as uniform pressure on a simply supported elastic plate. These solutions are sometimes exact (requiring only arithmetic evaluations of explicit mathematical expressions), but are often semianalytical (represented by infinite series, complex integrals, or asymptotic expansions). Difficulties can arise in computing any of these semianalytical solutions, especially infinite series. The modeler must ensure that when used for code verification, numerical error has been reduced to an acceptable level.

For problems that allow analytical solutions, whether exact or semianalytical, pass/fail criteria can be stated in terms of the following two types of comparison:

(a) the agreement between the observed order of accuracy and the formal order of accuracy of the numerical method
(b) the agreement of the converged numerical solution with the analytical solution using specified norms

When computational solutions are compared with analytical solutions, either the comparisons should be examined in the regions of interest or the error norms should be computed over the entire solution domain. The accuracy of each of the dependent variables or functionals of interest should be determined as part of the comparison.

5.1.1.2 Method of Manufactured Solutions (MMS). The MMS is a technique for developing a special type of analytical solution (refs. [14], [18]). To apply it, the modeler prescribes solution functions for the PDEs and finds the forcing functions that are consistent with the prescribed solution. That is, the prescribed solution functions are inserted into the PDEs, and the equations are rearranged such that all remaining terms in excess of the terms in the original PDEs are grouped into forcing functions or source terms. Initial conditions and boundary conditions are similarly derived, based on the prescribed solution on the boundary. For example, for the simply supported plate problem, one could prescribe a solution of displacements that requires a highly variable pressure distribution or even applied internal moments. If this pressure and moment "forcing function" can be derived, it can then be applied using a computational model for the plate, and the computed displacement field can be compared to the prescribed solution.

The advantages of the MMS are many. It can be applied to a wide variety of highly nonlinear problems. It can test a large number of numerical features in the code, such as the numerical method, the spatial-transformation technique for mesh generation, the mesh distribution technique, and the correctness of algorithm coding (ref. [14]). The MMS provides a clear assessment because, unless there are software errors, the computational results must agree with the solution used to derive the forcing function.

The MMS is not without its disadvantages. In any nontrivial application of this method, the algebra and calculus required to derive the forcing function can become very complex, and symbolic manipulation software may offer the only practical recourse. Using the MMS can also require special coding and compilation if the code does not admit separate externally applied nodal forces for every degree of freedom at every node, each with its own time history. While the MMS can efficiently highlight the presence of errors, it cannot point to the sources of these errors and cannot identify mistakes in algorithm efficiency (refs. [14], [18]).

5.1.1.3 Numerical Benchmark Solutions. When analytical solutions cannot be found or derived, the only other option for benchmark solutions is numerically derived ones. There are two distinct categories of highly accurate numerical benchmark solutions. One category consists of solutions in which the PDEs have been reduced by similarity transformations or other means to one or more ODEs that must be integrated numerically. The other category consists of solutions in which the PDEs have been solved directly by numerical methods. The accuracy of such numerical benchmark solutions has to be critically assessed to qualify them for use in code verification. For the numerical integration of ODEs, well-established standard methods are available for assessing accuracy. In the case of numerically integrated PDEs, no published solution can be considered a benchmark until the code used in producing that solution has been thoroughly verified and documented. In addition, comprehensive numerical error estimation must be reported. Credibility will be enhanced if independent investigators, preferably using different numerical approaches and computer software, produce multiple solutions that agree. Using multiple independent sources for the solutions will mitigate the risk of errors in the verification benchmark.

5.1.1.4 Consistency Tests. Consistency tests can be used to verify numerical algorithms. Global as well as local tests should be made for the conservation of mass, momentum, and energy (ref. [19]). An algorithm can satisfy the conservation laws exactly, or it can satisfy the laws in the limit of infinite resolution; this distinction should be considered when assessing the accuracy of an algorithm. Consistency tests can also be made that involve geometry (e.g., checking that the same numerical solution is obtained in different coordinate systems or determining whether specific symmetry features are preserved in the solution). Consistency tests should be considered complementary to the other types of algorithm tests described herein for numerical algorithm verification. If they can be devised, consistency tests are especially important because the failure of these tests indicates that there are unacceptable errors in the code.

5.1.2 Software Quality Engineering (SQE). The SQE part of code verification refers to procedures that provide evidence that the software implementation of the numerical algorithms is free of programming errors and implementation faults. Such errors most commonly reside in the source code, but occasionally flaws in the compiler introduce them. Evidence of error-free software from SQE is a necessary element of verification. SQE determines whether the software system is reliable and produces reliable results on specified computer hardware with a specified software environment (compilers, libraries). To optimize its influence on code verification, SQE should be planned and used during the development of the software product, not as a retrospective activity for a fielded software implementation (ref. [20]). However, feedback from users to developers is highly encouraged.

5.2 Calculation Verification

Calculation verification is applied to a computational model that is intended to predict any simulation results. Thus, each computational model developed in a validation hierarchy is subject to calculation verification. The goal of calculation verification is to estimate the numerical error associated with the discretization. In most cases, exercising the computational model with multiple meshes is required to estimate this error. Another source of error is mesh bias, wherein the arrangement of the elements can influence the results, especially if the mesh is coarse.

The two basic approaches for estimating the error in a numerical solution to a complex set of PDEs are a priori and a posteriori. A priori approaches use only information about the numerical algorithm that approximates the partial differential operators and the given initial and boundary conditions. A posteriori error estimation approaches use all of the a priori information plus the results from two or more numerical solutions to the same problem that have different mesh densities and/or different time steps (refs. [14], [21], [22]). The discussion here focuses on a posteriori error estimates because they can provide quantitative assessments of numerical error in practical cases of nonlinear PDEs.

5.2.1 A Posteriori Error Estimation. A posteriori error estimation has primarily been approached using either finite-element–based error estimation techniques (refs. [23], [24]) or multiple-mesh solutions combined with Richardson extrapolation and extensions thereof (ref. [14]).

Two fundamentally different types of finite-element–based discretization error estimators have been developed. The most commonly used are recovery methods, which involve post-processing of either solution gradients or nodal values in patches of neighboring elements. These provide direct error estimates only in the global energy norm; however, they provide ordered error estimates for specific field quantities of interest (i.e., the estimate improves with mesh refinement).

The second class of finite-element–based error estimators consists of residual-based methods. Like recovery methods, residual methods were originally formulated to provide error estimates in the global energy norm. Extension to error estimates in response quantities of interest, such as deflections or stresses, generally require additional solutions (ref. [25]).

Single-mesh finite-element–based error estimates, when applicable, offer a great advantage by reducing mesh-generation and computational effort. However, the estimates require that the convergence rate be assumed. Calculation of an observed convergence rate always requires the generation of multiple meshes. The single-mesh a posteriori methods are also important for finite element adaptivity, where both the spatial mesh density (known as h-adaptivity) and the order of the finite element scheme (known as p-adaptivity) can be adapted (refs. [23], [24]).

Standard Richardson extrapolation assumes that

(a) the observed order of accuracy (rate of convergence) is known

(b) two numerical solutions at different mesh resolutions have been computed

(c) both solutions are in the asymptotic convergence regime

To estimate a bound on the numerical error, the Richardson method then extrapolates to a more accurate value against which to compare the original solution. Various elaborations of Richardson extrapolation use three or more meshes to calculate an observed order of accuracy (refs. [6], [14]). The observed order of accuracy can be used to verify a theoretical order of accuracy, test whether the solution is in the asymptotic regime, and estimate a zero-mesh-size converged solution using extrapolation. A grid convergence index (GCI) based on Richardson extrapolation has been developed and advocated to assist in estimating bounds on the mesh convergence error (refs. [14], [26]). The GCI can convert error estimates that are obtained from any mesh-refinement ratio into an equivalent mesh-doubling estimate. More generally, the GCI produces an error-bound estimate through an empirically-based factor of safety applied to the Richardson error estimate (refs. [6], [14]).

5.2.2 Potential Limitations. The assumption of smoothness in solutions (i.e., the absence of singularities and discontinuities) underlies much of the theory of existing error estimation techniques and is quite demanding in estimating local errors in the solution domain; however, this assumption does not prevent the use of an empirical approach to error estimation based on observed convergence rates. Experience shows that an empirical approach is more dependable when more than three meshes are used with a least squares evaluation of observed convergence rates and when functionals rather than point values are considered.

Singularities and discontinuities commonly occur in solid mechanics; the crack tip singularity is an example. The difficulties of singularities and discontinuities are compounded in very complex conceptual models, where multiple space and time scales may be important and very strong nonlinearities may be present. Ideally, calculation verification should be able to confront these complexities. However, the "pollution" of particular regions of a calculation by the presence of singularities such as shock waves, geometrical singularities, or crack propagation is a subject of concern in error estimation (refs. [14], [24], [27]), and there is a lack of rigorous theory for guidance in these situations.

Another complexity in numerical error estimation is the coupling that can occur between numerical error and the spatial and temporal scales in certain types of physical models. Refining the mesh does not ensure that the physics modeled will remain unchanged as the mesh is resolved. For example, an insufficiently refined mesh in buckling problems will prevent the model from exhibiting higher modes of buckling. This observation regarding mesh refinement directly influences the accuracy and reliability of any type of a posteriori error estimation method, especially extrapolation methods.

5.3 Verification Documentation

Documentation must be an integral part of the verification process to facilitate reuse of the model. The documentation should explain the rationale and limitations of the code verification and calculation verification activities. It should include descriptions of the error estimation techniques employed, the results of consistency tests, and the analytical solutions, manufactured solutions, and numerical benchmark solutions used. SQE and SQA, configuration management, and acceptable computational systems should also be described.

6 VALIDATION

The activities described in this section are performed for each validation case in the validation hierarchy developed during preparation of the V&V plan.

Although the immediate goal of validation is to compare simulation results with experimental measurements, the strategic goal is to increase confidence in the predictive capability of a computational model for its intended use. This is accomplished by comparing computational predictions (simulation outputs) to observations (experimental outputs). Three prerequisites for validation are

(a) a clear definition of the model's intended use

(b) completed code verification and calculation verification activities conducted sufficiently so that the errors discovered through validation can be isolated from those errors discovered through verification

(c) quantified uncertainties in both the simulation outputs and the experimental outputs

The approach of validation is to measure the agreement between the simulation outputs from a computational model and the experimental outputs from appropriately designed and conducted experiments. These outputs should incorporate the experimental and modeling uncertainties in dimensions, materials, loads, and responses. In most cases, the assessment of the predictive capability of a computational model over the full range of its intended use cannot be based solely upon data already available at the beginning of the V&V program. Not only might existing data inadequately represent the intended use of the model, it may also have been used in model calibration during the development of the computational model. In such cases, new experiments and computational predictions are required. The challenge is to define and conduct a set of experiments that will provide a test of the model stringent enough that the decision maker will have adequate confidence to employ the model for predicting the validation case. If the model predicts the experimental outputs within the predetermined accuracy requirements, the model is considered validated for its intended use.

6.1 Validation Experiments

Validation experiments are performed to generate data for assessing the accuracy of the mathematical model via simulation outputs produced by the verified computational model. A validation experiment is a physical realization of a properly posed applied mathematics problem with initial conditions, boundary conditions, material properties, and external forces. To qualify as a validation experiment, the geometry of the object being tested (e.g., a component, subassembly, assembly, or full system), the initial conditions and the boundary conditions of the experiment, and all of the other model input parameters must be prescribed as completely and accurately as possible. Ideally, this thoroughness on the part of the experimenter will provide as many constraints as possible, requiring few assumptions on the part of the modeler. All of the applied loads, multiple response quantities, and changes in the boundary conditions should be measured, and uncertainties in the measurements should be reported.

6.1.1 Experiment Design. Generally, data from the literature are from experiments performed for other purposes and thus do not meet the requirements of a validation experiment. Experiments can have many purposes and are often focused on assessing component performance relative to safety criteria or exploring modes of system response. Consequently, the measurement set in many experiments may differ from the measurements needed for model validation. For example, a test may show that a component fails at a load higher than an acceptable threshold and thereby establish that the component is acceptable for use. However, the test may not have measured the deformation as the force was applied because that measurement was not needed for the purpose of the experiment. If both the component-failure measurement and the deformation measurement were necessary to validate a computational model, the test measuring only component failure could not be used for validation. Furthermore, predictions of experiments whose results are known prior to the validation effort are influenced, even if subconsciously, by modelers' assumptions, knowledge of the experimental results, and the selection of unmeasured quantities. For these reasons, it is usually necessary to perform experiments that are dedicated to model validation (refs. [3], [6]).

The modeler should have input regarding the design of the validation experiments. The experimenter and the modeler need to share an understanding of the responses that are difficult to measure or predict. The modeler needs to be certain that all the inputs (especially for constitutive models), boundary conditions, and imposed loads are being measured. The modeler should perform a parametric study with the verified model to determine model sensitivities that need to be investigated experimentally. In addition, pretest analyses should be conducted to uncover potential problems with the design of the experiment. However, credibility of the validation process will be greatly enhanced if the modeler does not know the test results before the prediction is complete, with the exception that the modeler must be provided material properties, applied loads, and initial and boundary conditions.

In summary, the validation experiments and measurement set should be designed to leave as few unknown parameters as possible. In the all-too-common case that some significant parameters are not measured, the modeler has to perform multiple calculations to compare with the experiments by varying the values of those parameters. The modeler cannot arbitrarily select a parameter value within its accepted range and base the validation comparison on that selection because doing so can result in either false validation or false invalidation. If all of the calculation results using a realistic range of the parameters are within the acceptable tolerance for validation, then validation may be claimed, even though the experiment had uncontrolled variables. But if the calculation results for a significant portion of the realistic parameter range lie outside this tolerance, validation cannot be claimed, and progress can only be made by the experimenter constraining the range of the unmeasured or poorly measured parameters.

6.1.2 Measurement Selection. Selection of the quantities to measure should be based primarily on the response quantities of interest. When possible, these quantities should be measured directly rather than derived from other measurements. For example, if strain is the quantity of interest, it is probably better to use a strain gage instead of multiple measurements of displacement. Similarly, if velocity can be measured directly, that approach is better than integrating a measurement of acceleration or differentiating a measurement of displacement. On the other hand, consistency of the test data is an important attribute that increases confidence in the data. While it is recommended to use direct measurements for the primary response quantity, data consistency can be established by supplementing these with corroborative measurements derived independently (e.g., measuring displacement or acceleration to corroborate measurements of velocity). Measurements of point quantities made in families that allow fields to be estimated are also useful; for example, a displacement field can be used to corroborate measurements of strain (ref. [28]).

Another reason that variables or locations in the model other than those specified in the validation requirements should be measured is that agreement between these measurements and the simulation results can contribute significantly to overall confidence in the model. Although some quantities may be of secondary importance, accurate calculations of these quantities provide evidence that the model accurately calculates the primary response for the right reason. For example, confidence in a model that matches the central deflection of a beam is greatly enhanced if it also matches the displacements or strains all along the length of the beam — even if central deflection is the only quantity of interest for the intended use. This can qualitatively or even quantitatively build confidence that the model can be used to make accurate predictions for problem specifications that are different from those included in model development and validation. Thus, validation experiments should produce a variety of data so that multiple aspects of the model can be assessed.

6.1.3 Sources of Error. It is important to calibrate the gages that will be used in validation experiments and to document their inaccuracies related to nonlinearity, repeatability, and hysteresis. Many things can influence the output of a gage. Pressure transducers, for example, should be calibrated in an environment similar to that of the validation experiment (e.g., at elevated temperature). If a transducer is sensitive to the environment and the environment changes significantly during a validation test, the transducer's sensitivity to the environment must already have been established (during previous calibration of the gage) so that the resulting data can be corrected to account for it (ref. [29]).

In addition, the experimenter needs to determine and account for effects such as the compliance or inertia of any test fixtures if these effects contribute to the measurement of displacement or force, respectively. For example, the mass of a piston in a hydraulic testing machine can affect the measurement of the force applied to the specimen and, if ignored, can contribute to lack of agreement between the simulation results and the experimental results. Reporting the details of operating, calibrating, and installing the gages used in an experiment helps the modeler understand the relationship between gage output and model output. It may even be necessary in some cases for the modeler to build a model that includes such parts as the test fixtures or measurement fixtures to accurately predict the measurements.

6.1.4 Repeated Measurements. For validation experiments, redundant measurements are needed to establish the precision (scatter) in the validation test results and thus improve the quantification of uncertainty in experimental measurements. One approach for obtaining redundant measurements is to repeat the test using different specimens. The test-to-test scatter would then have contributions from differences in specimens (initial conditions) or material properties, specimen installation (boundary conditions), gages, gage installation, and data acquisition. As an example, if bending tests were performed on several members of a set of beams and the responses measured with strain gages mounted on the tension and compression surfaces, not only would each beam be different, but each might be off center in the testing machine by differing amounts. In addition, the strain gages would have different scatter in location and orientation, and the signal-wire resistances would differ.

Another approach for obtaining redundant measurements is to repeat the test using the same specimen. This approach may be taken if the cost of testing is high or the availability of test specimens is limited. Of course, specimen-to-specimen response variability would not be obtained. Still another approach for obtaining redundant measurements is to place similar transducers at symmetrical locations (if the test has adequate symmetry) to assess scatter. The data from these transducers could also be used to determine whether the expected symmetry was indeed obtained.

6.2 Uncertainty Quantification in Experiments

In the UQ activity for experiments, the effects of measurement error, design tolerances, construction uncertainty, and other uncertainties are quantified, resulting in the experimental outputs. Although published experimental results often do not include an assessment of uncertainty, it is necessary to estimate and report the uncertainty in the measurements in validation experiments so that simulation results can be judged appropriately.

In experimental work, errors are usually classified as being either random (precision) or systematic (bias). An error is classified as random if it contributes to the scatter of the data in redundant measurements or repeat experiments at the same facility. Random errors are inherent to the experiment, produce nondeterministic effects, and cannot be reduced with additional testing, although they can be better quantified with additional testing. Sources of random error include dimensional tolerances on test parts, variability in assembly and measurement locations, variability of material properties, and mechanical equipment variances due to friction. Systematic errors can produce a bias in the experimental measurements that is difficult to detect and estimate. Sources of systematic error include transducer calibration error, data acquisition error, data reduction error, and test technique error (ref. [30]).

Either the experimenter or an independent reviewer must provide an uncertainty assessment of the results. The assessment should consider all sources of experimental uncertainty, whether the sources were measured or estimated. When possible, the uncertainties should take the form of mean values with standard deviations or distributions (ref. [3]). Even when statistics are not available, an estimate of experimental uncertainty based on previous experience or expert opinion is necessary before proceeding to comparisons with simulation outputs. A common pitfall is to neglect important contributions to modeling uncertainty, experimental uncertainty, or both, and then try to draw conclusions about predictive accuracy based on inadequate information. Improper or inappropriate inferences could thus be made about the accuracy of the computational model.

6.3 Model Accuracy Assessment

Following UQ of the experimental results that produced the experimental outputs, the final steps in validation consist of *(a)* comparing values of the metrics chosen to measure the agreement between simulation outputs and experimental outputs

(b) making an assessment of the accuracy of the computational model relative to the goals provided in the V&V plan for the model's intended use

Recall that a model accuracy assessment (see Figure 3.3-1) is made for each component, subassembly, assembly, and subsystem in every level of the validation hierarchy for which validation data are produced. The determination of the system-level model's accuracy is made after the hierarchy of validation experiments has been performed and the composite computational model has been validated through the various hierarchical tiers.

6.3.1 Validation Metrics. A validation metric provides a method by which the simulation outputs and the experimental outputs can be quantitatively compared. The metric result is compared to the accuracy requirements defined in the V&V plan to determine whether acceptable agreement has been achieved. Validation metrics should incorporate the uncertainties associated with the experimental outputs and the uncertainties associated with the simulation outputs (e.g., the input parameter uncertainties propagated through the computational model). When multiple (repeat) experiments have been performed, the mean and variance of the system response of interest can be quantified. Multiple metrics that quantify the difference between uncertain model and test outputs have been proposed for scalar quantities (e.g., refs. [31], [32], and [33]).

Which experimental and simulation outputs to compare should be carefully considered. The outputs of interest may be simple, such as tip deflection or quarter point strain, or more complex, such as a comparison of spatial or temporal distributions (e.g., strain as a function of distance, or velocity at a point as a function of time). Many aspects of validation metrics, including the comparison of spatial or temporal distributions, are active areas of research with no general, proven methodology for real-world problems.

Validation metrics can sometimes be devised to incorporate the uncertainties associated with the experimental and simulation outputs (e.g., the input parameter uncertainties propagated through the computational model). When multiple (repeat) experiments have been performed, the mean and variance of the system response of interest can be quantified. A metric for the special case of multiple experiments with no uncertainty in the simulation outputs has been proposed (refs. [34], [35], [36]). For the general case, where both the measurement and simulation are expressed as a mean with variance, some research has been performed (ref. [32]), but this and other aspects of validation metrics are still active areas of research.

6.3.2 Determination of Accuracy. It is possible the model will fulfill only a portion of the validation requirements. The accuracy may fall short of the requirements in general or for a certain portion of the intended use. For example, a 10% accuracy goal may be unmet, but 15% accuracy may be established. Alternately, the 10% accuracy may be met for loads under or over a given level or for all but a particular type, such as thermal. Assuming that the original criteria were properly established for the intended use, this implies that further model improvements are needed. In the meantime, the model may have utility on a limited basis (i.e., it may be validated to a lower standard than that specified in the V&V plan, or it may be partially validated). In such cases, the technical experts and decision makers have the shared burden of establishing partial acceptance criteria. They could establish a new and less-ambitious definition of the acceptable level of agreement for validation, or they could define the limitations of the model's use. Partial validation is not uncommon, which underscores that a verdict or claim of "validation" is never meaningful without reporting the accuracy criteria and the uncertainties in experiments and calculations.

Confidence in the model's predictions decreases as the conditions of application deviate from those used in the validation process. For example, a model of an engine block that has been developed to accurately predict the stresses on the cylinder surfaces may not adequately or accurately predict the stress near an internal cooling channel in the same model. Confidence in the model's output is limited to applications that are judged to be sufficiently similar to that for which validation was performed. Confident use for other purposes requires additional validation.

6.4 Validation Documentation

Documentation of the overall validation process and the specific validation activity (for each validation case in the hierarchy) conveys an understanding of the predictive capability of the model for its intended use and supports the conclusion about whether or not the model was successfully validated for the set-points of the validation experiments. The documentation also facilitates reuse of the knowledge base by enabling subsequent users to build upon the established validation activity, regardless of whether the model was successfully validated for its original intended use.

For each validation case, the validation documentation should build upon the documentation of the conceptual model and the documentation describing the verification process of the computational model. The resulting validation documentation is useful for answering essential questions such as "Are the approximations and uncertainties inherent in the