ASME VVUQ 10.2-2021

The Role of Uncertainty Quantification in Verification and Validation of Computational Solid Mechanics Models

AN INTERNATIONAL STANDARD

The American Society of Mechanical Engineers

ASME VVUQ 10.2-2021

The Role of Uncertainty Quantification in Verification and Validation of Computational Solid Mechanics Models

AN INTERNATIONAL STANDARD

The American Society of Mechanical Engineers

Two Park Avenue • New York, NY • 10016 USA

Date of Issuance: May 31, 2022

The next edition of this Standard is scheduled for publication in 2026.

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 VVUQ Committee may be published as Cases. Cases are published on the ASME website under the VVUQ 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 VVUQ 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 ensure that individuals from competent and concerned interests had an opportunity to participate. The proposed code or standard was made available for public review and comment, which provided 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 does ASME 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 representatives or persons 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 iv
Committee Roster v
Correspondence With the VVUQ Committee vi
Introduction viii

1 **Purpose and Scope** 1
2 **Background and Definitions** 1
3 **Uncertainty Quantification in Modeling and Simulation** 9
4 **Uncertainty Quantification in Validation Experiments** 14
5 **Uncertainty Quantification in Model Validation Assessment** 16
6 **Uncertainty Quantification in Revisions to Model and Experiment** 20
7 **Uncertainties in Hierarchical Models** 21
8 **References** 22

Figures
1.1-1 Verification, Validation, and Uncertainty Quantification Process 2
2.2-1 Illustration of a Tapered Cantilever Beam 3
2.6-1 Illustration of Aleatory and Epistemic Uncertainty 7
2.8-1 Variation of Standard Residual Error With Number of Input Variables Included in Empirical Model Formulation 9
3.1.2-1 Defining a Random Variable Using Data 12
3.2-1 Illustration of the Uncertainty Propagation Process 13
5.2.1-1 Extension of the Mean Metric: The Difference Between Y_{mod} and Y_{exp} 17
5.2.2-1 Illustration of the Area Metric 18
5.2.3-1 PDF and CDF of the Error Metric 19
7-1 Illustration of a Model Hierarchy 22

Table
2.3-1 Examples of Sources of Error and Uncertainty 5

FOREWORD

Verification, validation, and uncertainty quantification (VVUQ) of engineering simulation models is currently of great interest to government, industry, and academia, all of which rely on computational modeling in the design and analysis of engineered structures. This has challenged government, industry, and academia to develop credible models for design and performance evaluation via computational simulation. In response, The American Society of Mechanical Engineers (ASME) formed a subcommittee on verification and validation (V&V) in computational solid mechanics (CSM), known as ASME V&V 10. The V&V 10 Subcommittee was redesignated as VVUQ 10 in 2021 since both V&V and uncertainty quantification (UQ) have been specifically identified as critical technologies needed for the advancement of computational mechanics.

The growing reliance on CSM models for decision-making in government, industry, and academia demands that greater attention be given to the quantification of uncertainties associated with these models. UQ is a critical component in the evaluation and communication of both computational and experimental results in the process of reporting simulation results. UQ is foundational in the development and assessment of CSM models and their predictive capability.

ASME V&V 10-2006 (ref. [1]), Guide for Verification and Validation of Computational Solid Mechanics Models, was recently revised to ASME V&V 10-2019 (ref. [2]), Standard for Verification and Validation of Computational Solid Mechanics Models. Both editions address an important need for a common language and process definition for VVUQ at the level appropriate for CSM model developers as well as their managers and customers. ASME V&V 10 is intended as an overview standard to be followed by detailed publications covering select topics and applications. ASME V&V 10.1 (ref. [3]) provides examples to illustrate many key VVUQ concepts. The overall purpose of this Standard is to expand on UQ, emphasized in ASME V&V 10.

This Standard describes the role of UQ in modeling/simulation and experimentation. UQ in modeling and simulation includes consideration of model form uncertainties, numerical solution uncertainties, model input uncertainties, and uncertainties in model-basis data. In addition, propagation of uncertainties is an integral part of UQ in modeling and simulation. UQ plays an important role in experimentation; therefore, key considerations in planning validation experiments are discussed, since these experiments are specifically planned and performed to assess the predictive capability of a computational model. A brief discussion of UQ in hierarchical CSM models is provided, as well as an overview of the role of UQ in revisions to either the computational model or the validation experiment.

This Standard also addresses the role of UQ in model validation assessment, illustrated by several examples considering different validation metrics that include uncertainties. Because validation metrics incorporate uncertainties in both experimental measurements and model simulations, some consideration of uncertainties is required in establishing the corresponding validation requirements.

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 VVUQ 10.2-2021 was approved by the VVUQ Standards Committee on August 27, 2021, and was approved and adopted by the American National Standards Institute on December 13, 2021.

ASME VVUQ COMMITTEE
Verification, Validation, and Uncertainty Quantification in Computational Modeling and Simulation

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

STANDARDS COMMITTEE OFFICERS

D. M. Moorcroft, *Chair*
B. H. Thacker, *Vice Chair*
M. Pagano, *Secretary*

STANDARDS COMMITTEE PERSONNEL

M. D. Benedict, U.S. Air Force Research Laboratory
J. Bischoff, Zimmer Biomet
D. Cheng, Fluor Corp.
S. W. Doebling, Los Alamos National Laboratory
K. Dowding, Sandia National Laboratories
L. Eca, Instituto Superior Tecnico
C. J. Freitas, Southwest Research Institute
Y. Hassan, Texas A&M University
M. Horner, ANSYS, Inc.
J. S. Kaizer, U.S. Nuclear Regulatory Commission
K. E. Lewis, Lawrence Livermore National Laboratory
S. Mahadevan, Vanderbilt University
D. M. Moorcroft, Federal Aviation Administration
M. Pagano, The American Society of Mechanical Engineers
S. Rachuri, U.S. Department of Energy
R. Schultz, Consultant
V. Sharma, Engineering Systems, Inc.
B. H. Thacker, Southwest Research Institute
T. M. Morrison, *Contributing Member,* U.S. Food and Drug Administration

VVUQ 10 SUBCOMMITTEE — VERIFICATION, VALIDATION, AND UNCERTAINTY QUANTIFICATION IN COMPUTATIONAL SOLID MECHANICS

D. M. Moorcroft, *Chair,* Federal Aviation Administration
S. L. Kieweg, *Vice Chair,* Sandia National Laboratories
M. Pagano, *Secretary,* The American Society of Mechanical Engineers
M. C. Anderson, National Nuclear Security Administration
S. Atamturktur, Pennsylvania State University
G. Banyay, Pennsylvania State University
M. D. Benedict, U.S. Air Force Research Laboratory
J. Bodner, Medtronic Corp.
J. Budzien, Los Alamos National Laboratory
J. V. Cox, Sandia National Laboratories
R. M. Ferencz, Lawrence Livermore National Laboratory
L. Gutkin, Kinectrics
J. F. Imbert, JFI Consult
A. Koskelo, Los Alamos National Laboratory
D. Lancaster, Pratt and Whitney
I. Lopez, Lawrence Livermore National Laboratory
W. L. Oberkampf, William L. Oberkampf Consulting
C. F. Popelar, Southwest Research Institute
C. Robeck, Advanced Micro Devices
B. H. Thacker, Southwest Research Institute
G. Orient, *Alternate,* Sandia National Laboratories
S. W. Doebling, *Contributing Member,* Los Alamos National Laboratory
W. Doeland, *Contributing Member,* European Union Aviation Safety Agency
T. K. Hasselman, *Contributing Member,* Consultant
H. U. Mair, *Contributing Member,* Johns Hopkins University, Applied Physics Laboratory
S.-R. Hsieh, *Contributing Member,* Lawrence Livermore National Laboratory
T. L. Paez, *Contributing Member,* Thomas Paez Consulting
X. Qu, *Contributing Member,* The Aerospace Corp.
C. P. Rogers, *Contributing Member*, Crea Consultants, Ltd.
G. Sagals, *Contributing Member,* Canadian Nuclear Safety Commission
J. F. Schultze, *Contributing Member,* Los Alamos National Laboratory
K. Teferra, *Contributing Member*, Consultant

CORRESPONDENCE WITH THE VVUQ 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, VVUQ 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.

This Standard is always open for comment, and the Committee welcomes proposals for revisions. 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 VVUQ 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 VVUQ 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 VVUQ 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 VVUQ 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 VVUQ Standards Committee. Future Committee meeting dates and locations can be found on the Committee Page at http://go.asme.org/VnVcommittee.

INTRODUCTION

Model verification, validation, and uncertainty quantification (VVUQ) are the primary methods for quantifying and building credibility in mathematical/computational models. Verification is the process of determining that a computational model accurately represents the underlying mathematical model and its solution. Validation is the process of determining the degree to which a model is an accurate representation of the validation experiments from the perspective of the intended uses of the model. Both verification and validation accumulate evidence of model correctness and accuracy for a specific application of interest. Model VVUQ* cannot prove that the model is correct and accurate for all possible scenarios; instead, VVUQ accumulates evidence of whether the model is sufficiently accurate for its intended uses. The expected outcome of the VVUQ process is to inform the decision-maker of the credibility of the model for the intended uses of the model.

Uncertainty quantification (UQ) in the context of VVUQ is the mathematical assessment of uncertainties in model simulation results and experimental results. Therefore, the goal is to quantify the uncertainties in both simulation and experimental results so that the model accuracy can be assessed, and the predictive capability of the model can be established quantitatively. As described in this Standard, UQ is also important in other related activities, including model development, parameter estimation, model calibration, experimental design, and sensitivity analysis.

* For the purposes of this Standard, "model VVUQ" will be referred to as "VVUQ."

THE ROLE OF UNCERTAINTY QUANTIFICATION IN VERIFICATION AND VALIDATION OF COMPUTATIONAL SOLID MECHANICS MODELS

1 PURPOSE AND SCOPE

1.1 Purpose and Motivation

The purpose of this Standard is to expand upon the important role of uncertainty quantification (UQ) in verification, validation, and uncertainty quantification (VVUQ), as outlined in Figure 1.1-1. UQ plays an important part in each of the "Modeling and Simulation" and "Physical Experimentation" branches illustrated in the figure, ultimately quantifying the uncertainties in the "Simulation Results" and "Experimental Results" generating the "Simulation Outputs" and "Experimental Outputs." A detailed description of this figure is provided in ASME V&V 10-2019.

Consistent with the purpose of ASME V&V 10-2019, the motivation for developing ASME VVUQ 10.2 is the need for a common language and process of UQ in computational solid mechanics (CSM), particularly as it may relate to how model developers perform UQ as well as how they subsequently communicate results, conclusions, and recommendations to a decision-maker. A decision-maker may be any individual or representative body, such as a review panel, deemed responsible for determining if a model is acceptable for its intended uses. The decision-maker may also be a customer relying on model predictions to inform a decision.

1.2 Objectives and Scope

1.2.1 Objectives. The objectives of this Standard are to

(a) define and clarify the role of UQ as part of the VVUQ process
(b) provide guidance for the use of UQ in VVUQ activities
(c) acknowledge the importance of UQ in decision-making

1.2.2 Scope. The scope of this Standard includes the following:

(a) sources and types of uncertainty and how they can be treated in the VVUQ process (section 2)
(b) quantification and propagation of uncertainties (section 3)
(c) uncertainties in validation experiments (section 4)
(d) uncertainties in model validation assessment (section 5)
(e) revisions to the model and experiments (section 6)
(f) uncertainties in hierarchical models (section 7)

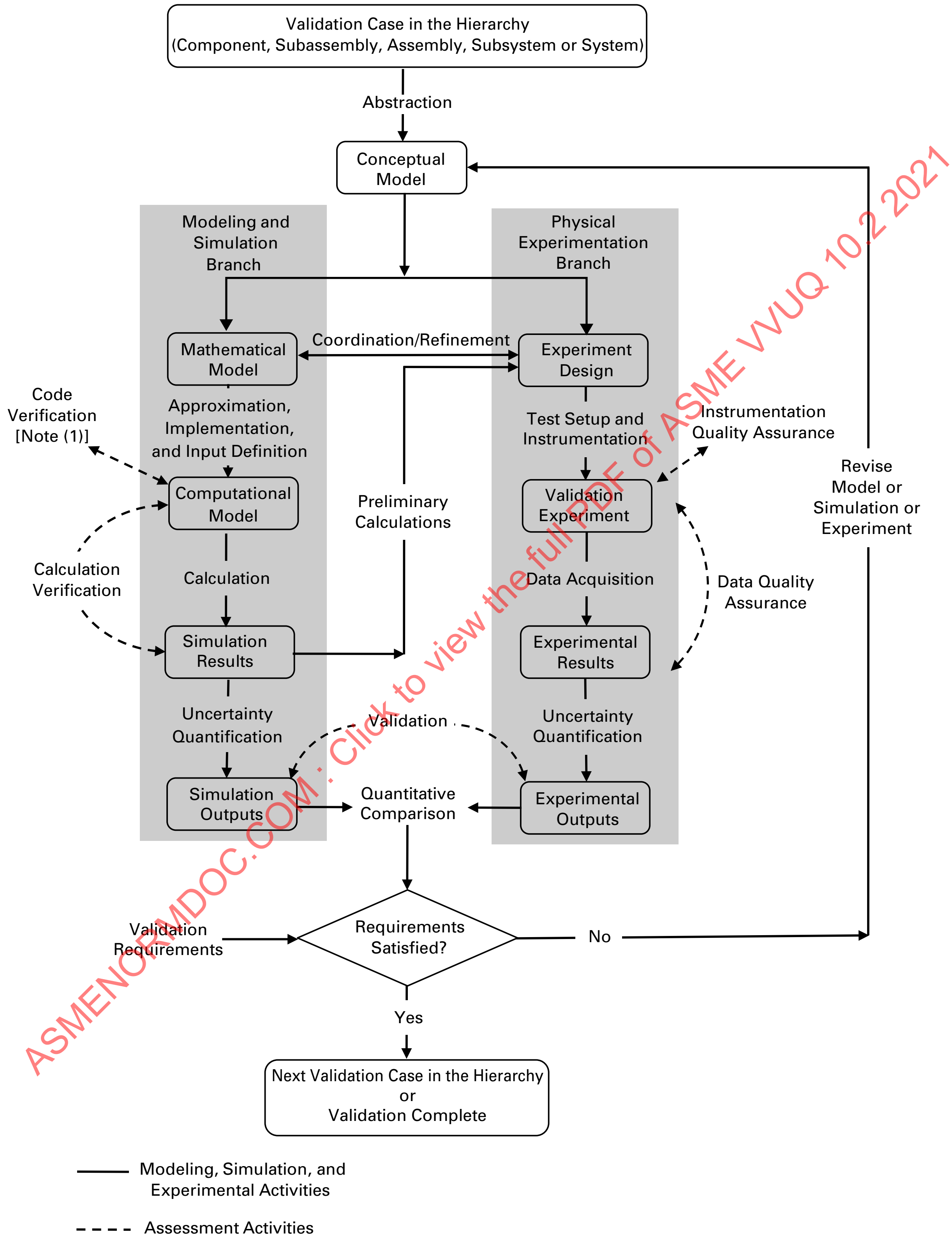
2 BACKGROUND AND DEFINITIONS

2.1 Mathematical Models

Models are idealized representations of the physical phenomena of interest. This Standard distinguishes between two types of mathematical models: empirical models and physics-based models. Examples of both types of mathematical models are provided in para. 2.2. These are defined as follows:

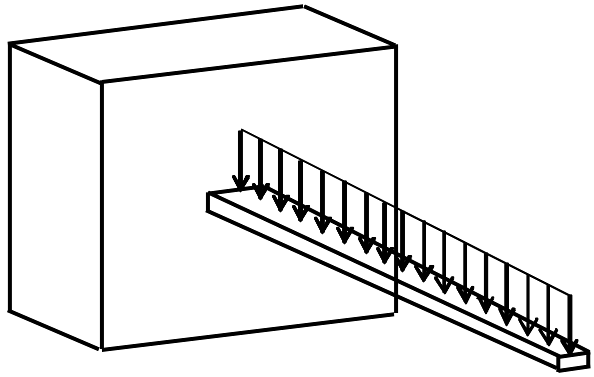
empirical model: a mathematical model whose functional framework is based primarily on observation and experiment. In CSM, empirical models are typically expressed in closed-form algebraic relations, which may still encompass some level of our conceptual or theoretical understanding of the physical phenomenon of interest. Empirical models may be simple statistical models (e.g., regression analysis) or more complex (e.g., based on machine learning or generalized polynomial expansions). Experimental data used to develop the empirical model are hereafter referred to as *model-basis data*. In the CSM community, this type of data is also referred to as *calibration data*.

Figure 1.1-1
Verification, Validation, and Uncertainty Quantification Process



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

Figure 2.2-1
Illustration of a Tapered Cantilever Beam



physics-based model: a mathematical model whose functional framework is based on concepts and theories capturing the physical phenomena of interest. Such models are alternatively referred to as scientific or mechanistic, and in CSM these models are often expressed by differential or integral equations or systems of equations. Experimental or observational information plays a secondary role in the development of physics-based models.

2.2 Variables and Parameters

A mathematical model of any type relates an output variable, y, to at least one input variable, x:

$$y = f(x) \tag{2-1}$$

where $f(\cdot)$ denotes the functional form of the mathematical model, hereafter referred to simply as the *model form*. Multiple input variables are common in practical CSM applications. It is noted that the variables are model-specific, and an input variable in one model may be an output variable in another model.

The model form $f(\cdot)$ itself includes additional quantities used to relate the output variable to the input variable(s) by means of mathematical operations. These quantities are referred to as *model parameters*. While some parameters may be physical or mathematical constants, the model parameters are typically quantities subject to variation, and therefore are generally uncertain. When a nonconstant parameter in a given model has physical meaning (para. 2.1), it may be an input variable or an output variable in another model.

In physics-based models, the model parameters typically possess definitive physical meaning. In empirical models, the model parameters may or may not have physical meaning. This distinction is very common and leads to differences in the process of uncertainty quantification due to the differences in the primary sources of uncertainty associated with the two types of models, as further discussed in para. 3.1.

Examples of variables and parameters in a physics-based model and in an empirical model follow:

(a) Example of Physics-Based Model. Consider the elastic deformation of a tapered cantilever beam shown in Figure 2.2-1. The vertical deflection, w, of the beam varies with the distance, x, from the supported end of the beam and the beam length, L, according to the following model derived from the classical beam theory (ref. [4]):

$$w(x, L) = \frac{qL^4}{EI_0} W\left(f_r, \alpha, \frac{x}{L}\right) \tag{2-2}$$

where

E = modulus of elasticity
f_r = support rotational flexibility
I_0 = area moment of inertia (at x = 0)
L = beam length
q = a distributed static load
w = vertical deflection
$W(\cdot)$ = function representing the dependence on f_r, α, and the ratio x/L
x = distance along beam from supported end
α = beam taper factor

In the physics-based model given by eq. (2-2), w is the output variable, x and L are the input variables, and the remaining five inputs are the parameters associated with this specific model form. It is noted that the input variables x and L are identified by adding them as arguments to the left-hand side of eq. (2-2).

If the interest is in the variation of the beam deflection w with the distance x from the supported end and the static load q, eq. (2-2) may be rewritten as

$$w(x,\ q) = \frac{qL^4}{EI_0} W\left(f_r,\ \alpha,\ \frac{x}{L}\right) \tag{2-3}$$

where w is the output variable, x and q are the input variables, and the remaining five quantities, including the beam length L, are the model parameters. Such reformulation is possible because both L and q have definitive meaning both within and outside of the specific model form used in this example.

(b) Example of Empirical Model. Consider the material resistance to failure due to fatigue that is characterized on the basis of experimental data obtained at different values, commonly referred to as levels, of alternating applied stress. The number of fatigue cycles to failure, N, is typically represented as a power-law function of the alternating applied stress, S

$$N(S) = b_0 \left(\frac{S}{S_0}\right)^{b_1} \tag{2-4}$$

where

b_0, b_1 = parameters estimated from experimental data
N = number of fatigue cycles to failure
S = alternating applied stress
S_0 = normalizing constant

In the empirical model given by eq. (2-4), by the definitions above, N is the output variable, S is the input variable, and S_0, b_0, b_1 are the parameters associated with this specific model form. The parameter S_0 is a normalizing constant, and the parameters b_0 and b_1 are estimated from experimental data. As in the example of a physics-based model in (a), the input variable S is identified by adding it as an argument to the left-hand side of eq. (2-4). In the current example, however, the parameters b_0 and b_1 do not possess definitive physical meaning and are only meaningful within the specific model form used in this example. Therefore, neither b_0 nor b_1 can be viewed as an input variable. On the other hand, the input variable S possesses definitive physical meaning both within and outside this model. Input variables such as S are also commonly referred to as explanatory variables. It is noted that the input variable S can take an arbitrary value within its range in the model application domain, whereas the model parameters b_0 and b_1 are represented by their best estimates and uncertainties.

The power-law functional form of S in this empirical model does not originate from understanding the physical mechanism of failure due to fatigue. Rather, it is selected for the following reasons:

(1) good track record with applicability to many structural materials
(2) versatility of trends obtained
(3) ease of linearization with respect to estimated model parameters, which substantially simplifies the process of model development

Indeed, taking logarithms of both sides of eq. (2-4) results in

$$\ln[N(S)] = \ln(b_0) + b_1 \ln(S/S_0) \tag{2-5}$$

and the formulation given by eq. (2-5) is then easily fit by linear regression to a set of experimental data $\{S_j, N_j\}$, where N_j is a value of N obtained experimentally at the jth level of the alternating stress S. The residual errors associated with such linear regression are discussed in para. 2.8.

2.3 Errors and Uncertainties

In this paragraph, error and uncertainty are defined and classified, with further discussion provided in para. 2.4.

(a) error: quantitative difference between a measured or calculated value and the referent or true value.

(b) uncertainty: lack of certainty due to inherent randomness and/or insufficient knowledge.

(1) aleatory uncertainty: uncertainty due to inherent randomness (irreducible).

(2) epistemic uncertainty: uncertainty due to lack of knowledge (reducible).

(-a) recognized: lack-of-knowledge uncertainty that is consciously recognized.

(-b) unrecognized: lack-of-knowledge uncertainty that is not known to exist.

Table 2.3-1
Examples of Sources of Error and Uncertainty

Source	Error/ Uncertainty	Type	Examples
Model	Error	Numerical	Errors associated with discretization, lack of mesh convergence, time stepping, numerical integration/differentiation, contact algorithms, hourglass controls, etc.
		Parametric	Incorrect parameter assertion
		Model form	Errors associated with idealized representation of true physics; geometric simplifications such as 2D plane strain/stress idealization; ignoring physical realities such as temperature effects
	Uncertainty	Epistemic	Unrecognized loading condition or failure mode; use of approximate probability models; assumption of probability models; small sample size
		Aleatory	Probability distributions for model parameters
Experimental setup	Error	Setup	Incorrect or incomplete experimental planning or execution
		Instrumentation	Exceeding physical limitations of experimental instruments; improper calibration
		Human	Misplacement of sensors/instruments; incorrect readings or transcriptions
	Uncertainty	Epistemic	Imperfect knowledge of loads/excitations and/or environmental conditions; measurement bias; insufficient number of experimental replications
		Aleatory	Variability in material properties, load/excitations, environmental conditions, and measurements
As-built test article	Error	Design	Misspecification of material requirements
		Construction	Improper or inferior materials; deviations from design specifications (as-built versus as-designed); varying bolt torque during assembly of a component
	Uncertainty	Epistemic	Unknown design tolerances; lack of knowledge about connectivity between components, method of construction, or boundary conditions; unknown material properties
		Aleatory	Load/excitations, environmental conditions

The true value of a quantity is known only in specific situations, e.g., when the value for a physical constant is defined by convention, such as the gravitational constant, or by international agreement in metrology.

The true value of any measurement in nature is never known exactly; it is typically only known to within an estimated uncertainty. Consequently, whether the true value is known or unknown, errors can be modeled as uncertainties. In most instances, this uncertainty would be epistemic in nature because it is theoretically reducible.

UQ in the context of VVUQ is the mathematical assessment of uncertainties in model simulation results and experimental results. Table 2.3-1 provides a breakdown of various sources, the different errors or uncertainties, and the types of each, with examples. This table only illustrates common classifications. In many cases, different uncertainty types can be modeled in different ways depending on the application problem, analyst preference, organizational culture, etc.

2.4 Aleatory and Epistemic Uncertainties

Aleatory uncertainty is also referred to as stochastic uncertainty, statistical uncertainty, irreducible uncertainty, variability, and inherent uncertainty. The fundamental nature of aleatory uncertainty lies in randomness, i.e., that of a stochastic process. Aleatory uncertainty commonly occurs in parameters that describe a system of interest. These may be, e.g., the stiffness (Young's modulus) of a material, the mass and geometric properties of a component or subsystem, or the stiffness in a bolted or riveted joint.

Epistemic uncertainty is also referred to as reducible uncertainty, knowledge uncertainty, and subjective uncertainty. The fundamental source of epistemic uncertainty is incomplete information or knowledge of any type that is related to modeling the system of interest, the environment the system is exposed to, and any approximations made during the formulation of the model. Epistemic uncertainty is associated with the modeler or observer, whereas aleatory uncertainty is associated with the system being modeled or observed.

Incomplete knowledge of the aleatory uncertainty results in a mixture of aleatory and epistemic uncertainty. Such an example is limited sampling. Including more samples reduces the epistemic uncertainty, but not necessarily the aleatory uncertainty. In fact, more samples may increase the aleatory uncertainty by improving knowledge of the true distributions.

Epistemic uncertainties are classified as recognized or unrecognized. A recognized epistemic uncertainty is one that acknowledges a limited state of knowledge exists concerning some aspect of modeling, and an attempt is made to quantify this limited knowledge. In this situation, when little or no experimental data are available for input quantities, it is common to resort to "expert" opinions from people knowledgeable of the system and environment. Some sources of recognized epistemic uncertainty are limited knowledge of the material properties of the system, boundary conditions, initial conditions, and system excitation.

Another important example of a recognized epistemic uncertainty is model-form uncertainty. This is uncertainty due to assumptions and approximations made in the formulation of the mathematical model of the system and environment. Generally, increasing the fidelity of the physics embodied within a model decreases the model-form uncertainty. However, higher fidelity also increases complexity, which in turn typically increases the number of parameters, computational burden, and the likelihood of errors.

An unrecognized epistemic uncertainty is commonly referred to as a blind uncertainty or an "unknown unknown." Common sources of unrecognized epistemic uncertainty are (unrecognized) human mistakes, misuse (either unintentional or intentional) of the system, mistakes in judgment concerning modeling of the system or environment, and undiscovered or misunderstood physical processes/mechanisms. Additional examples include unidentified programming mistakes in the simulation software, mistakes in the preparation of input data, and mistakes in recording or processing experimental data used for validation.

2.5 Deterministic and Nondeterministic Quantities

In the context of uncertainty quantification, both variables and parameters in a given model may be referred to as either deterministic or nondeterministic. These terms are defined as follows:

deterministic quantity: a variable or a parameter that is not considered uncertain and therefore can be assigned a fixed value.

nondeterministic quantity: a variable or a parameter that is considered uncertain and therefore cannot be assigned a fixed value.

The uncertainties associated with nondeterministic variables and parameters may be aleatory, epistemic, or a combination of the two, as defined in para. 2.4, and may be treated using probabilistic or nonprobabilistic methods, as discussed in para. 2.6.

With reference to the physics-based model represented by eq. (2-3), the input variable x and any of the model parameters are treated as either deterministic or nondeterministic quantities depending on whether or not the uncertainty associated with each quantity is of interest, or is deemed to have a significant effect on the output variable w. Given any uncertainty in the input variable x or in the model parameters, the output variable w is uncertain.

With reference to the empirical model represented by eq. (2-4), the input variable S may be treated either deterministically or nondeterministically depending on whether the uncertainty associated with S is of interest or is deemed to have a significant effect on the output variable N. The normalizing model parameter S_0 is assigned an arbitrary fixed value, i.e., treated deterministically. However, the model parameters b_0 and b_1 are estimated from experimental data. In engineering applications of this model, the best estimate values of b_0 and b_1 are commonly used to determine the variation of the output variable N with the input variable S. However, the nondeterministic nature of b_0 and b_1 is accounted for as described in para. 3.1.2.

2.6 Probabilistic and Nonprobabilistic Methods

Probabilistic methods treat uncertain variables and parameters of a mathematical model as random quantities described by probability distributions. These terms are defined as follows:

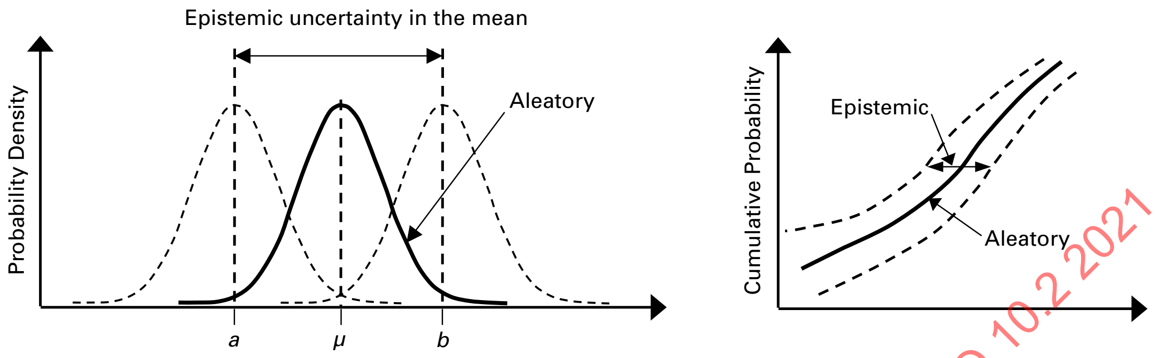
probability distribution: a mathematical relation describing how the probability associated with a variable or parameter varies with the value of the variable or parameter.

random quantity: a variable or parameter that is subject to inherent variations usually described by an associated probability distribution.

A probability distribution is typically expressed in the form of a probability density function (PDF) or a cumulative distribution function (CDF). A simple example is a normally distributed random variable,[1] X, with a PDF given by ref. [5]:

[1] In this Standard, the term "random variable" is used for both "random variable" and "random parameter" and will be the convention used.

Figure 2.6-1
Illustration of Aleatory and Epistemic Uncertainty



(a) Probability Density Function **(b) Cumulative Distribution Function**

GENERAL NOTE: The dashed lines represent other possible functions because of uncertainty in the parameters due to the use of limited sampling data (for example).

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{2-6}$$

where x is a specific value of X.

Examples of a PDF and a CDF for a normally distributed random variable are shown in Figure 2.6-1. Generally, the probability distribution (PDF or CDF) has its own parameters, such as the mean, μ, and standard deviation, σ, in eq. (2-6), which are typically estimated from data.

Two illustrations of aleatory and epistemic uncertainty are shown in Figure 2.6-1. In the figure, aleatory uncertainty is characterized by a PDF [illustration (a)] or alternatively a CDF [illustration (b)]. The PDF and CDF are alternative ways to describe the same uncertainty, and for simplicity the term "probability distribution" is used herein to refer to either representation.

A probabilistic representation is well suited for aleatory uncertainties (see para. 2.4) where an appropriate probability distribution can be constructed or assumed. In some cases, a PDF may also be used to represent epistemic uncertainties. For example, in Figure 2.6-1, a uniform distribution over an interval [a, b] may be selected to represent the epistemic uncertainty in the mean, μ, of the normal distribution used to represent the aleatory uncertainty. In other cases, a probability distribution is not appropriate for representing epistemic uncertainties. In such cases, the uncertainties could be treated using other mathematical theories such as intervals, probability bounds, or evidence theory (refs. [6], [7], [8]). For the nonprobabilistic representation of epistemic uncertainties, this Standard is confined to cases where uncertainty bounds can be reasonably established, enabling the use of interval methods. In Figure 2.6-1, for example, a uniform PDF mathematically assigns a constant probability to all values of μ in the interval in [a, b]. In contrast, simply assigning the interval [a, b] to the mean, μ, makes no assertion about the relative likelihood of any specific value of μ in the interval [a, b].

2.7 Sensitivity Analysis

Sensitivity analysis (SA) (ref. [9]) is performed to quantify the effects of changes in model inputs (e.g., material parameters, boundary conditions) on model outputs, i.e., response quantities (RQ).

Quantifying the relative importance of the model inputs can

(a) provide insight into the planning and justification of validation experiments
(b) facilitate model development decisions
(c) inform calibration activities

SA is often performed in conjunction with UQ for two reasons. First, SA can suggest which inputs have negligible impact on RQs and thus could be fixed during uncertainty propagation studies, resulting in reduced computational expense. Second, the characterization of input uncertainty can influence the results of SA.

There are two basic types of SA: local and global. Local SA involves perturbing model inputs, typically one at a time, around some nominal value to quantify the rates of change in model output(s) due to a change in a model input. Local SA does not consider the probability distribution of the model input. This deterministic approach can be problematic because a specific, local (nominal) value must be used. Most models of interest are also nonlinear, making local sensitivities of limited value except at or close to the specific point at which they were computed. However, the widespread use of local SA is due primarily to simplicity and efficiency.

Global SA is used to study how the variation in a model output can be apportioned to the various sources of variation in the model inputs (ref. [9]). These techniques are referred to as global because they consider the full variation in inputs, as compared to local SA methods that are used at a specific point. The two most common outcomes from a global SA are the main effect and total effect sensitivity indices. For a given model input, the main effect sensitivity index quantifies how much the output variance is reduced if the input variance is zero. The total effect sensitivity index is the main effect plus all interactions involving that input. Interaction effects are then quantified by taking the difference between the total and main effect indices.

Depending on the type of global SA performed, e.g., factorial analysis (ref. [10]) or analysis of variance (ANOVA) (ref. [11]), and the number of variates, global SA may require more computations than local SA. The number of simulation evaluations for a local SA scales with the number of nondeterministic inputs, M. On the other hand, variance-based decomposition (a typical method used to compute global sensitivity indices) may require $N(M + 2)$ where N is the user-specified number of sample evaluations. Ideally, the number of sample evaluations should be large enough that increased evaluations do not further change the global SA results.

2.8 Residual Errors and Residual Uncertainty

Consider the empirical model for material resistance to failure due to fatigue defined in para. 2.2. As discussed, such a model may be easily fit by single-variable linear regression to a set of experimental values of N obtained at different levels of S, if the model formulation is transformed as shown by eq. (2-5). In the process of its development, the linear regression model is typically represented as follows (ref. [11]):

$$\ln(N_j) = \ln(b_0) + b_1 \ln(S_j/S_0) + \varepsilon_j \tag{2-7}$$

where

$\{S_j, N_j\}$ = the model-basis set of data points indexed by j

$\{\mathcal{E}_j\}$ = associated vector of residual errors, i.e., differences between the observed values of N_j and the fitted values of N_j

The distribution of $\mathcal{E}_j$ values has the mean of zero, and its standard deviation is commonly referred to as the *standard residual error*, or simply as the *standard error*.

When the regression model developed as described is used to simulate N as a function of input variables, the vector of residual errors, $\{\mathcal{E}_j\}$, is replaced by a random quantity denoted as $U_{\mathcal{E}}$. For a single input variable, the alternating stress, S, the model is then formulated as follows:

$$N(S) = b_0 \left(\frac{S}{S_0} \right)^{b_1} \exp(U_{\mathcal{E}}) \tag{2-8}$$

$$\ln[N(S)] = \ln(b_0) + b_1 \ln(S/S_0) + U_{\mathcal{E}} \tag{2-9}$$

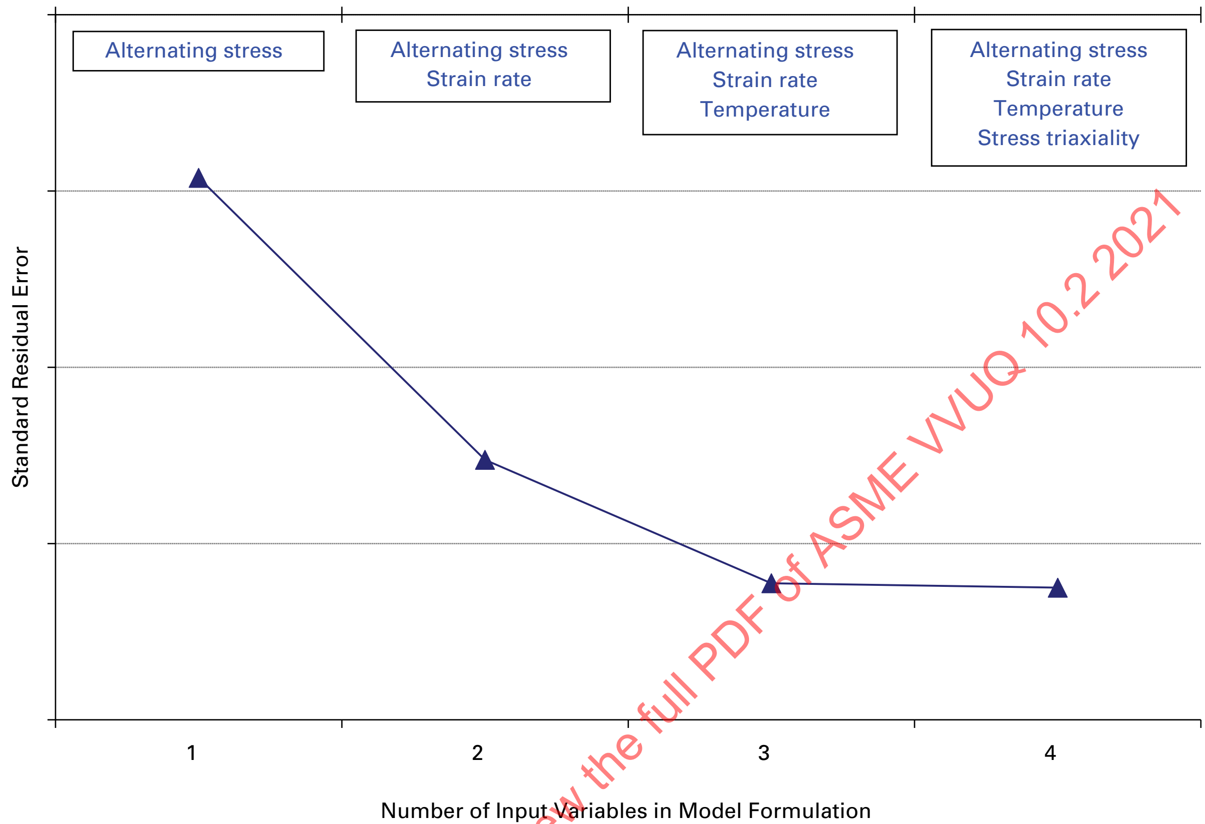
where

$U_{\mathcal{E}}$ = random quantity representing residual uncertainty

The formulations given by eqs. (2-8) and (2-9) are equivalent. The uncertainty $U_{\mathcal{E}}$ has a mean of zero and standard deviation proportional, but not equal, to the standard error. This is further discussed in para. 3.1.2.

This model as described has a single input variable, the alternating stress, S. It is quite common in engineering practice for the model-basis data to be associated with multiple input variables having more than one level. In this example, the model-basis data may involve different levels of strain rate, temperature, and stress triaxiality (due to the use of different specimen types in different testing rigs). Having the alternating stress as a single input variable in the model assumes that

Figure 2.8-1
Variation of Standard Residual Error With Number of Input Variables Included in Empirical Model Formulation



none of the other input variables affects the material resistance to failure due to fatigue to an extent that would warrant their incorporation into the model, and this may or may not be adequate.

A simple way of investigating the adequacy of this assumption is to incorporate an additive term into eq. (2-9) that is associated with each of the other input variables and to examine the results of multivariable linear regression analysis. Example results of such analysis are shown in Figure 2.8-1, representing the variation of the standard residual error with the number of input variables included in the model formulation. It is evident that considering the effect of both the strain rate and the temperature is associated with substantial reduction in the standard residual error, whereas the addition of the stress triaxiality to the model formulation does not result in any further decrease in this error.

Note that the standard residual error is not reduced to zero by incorporating additional input variables into the model formulation, i.e., improving the model quality this way eventually reaches a limit, which is determined by the inherent randomness in the model-basis data. The residual uncertainty has multiple sources, both aleatory and epistemic, and is therefore a mixed uncertainty. The sources of the residual uncertainty are further discussed in para. 3.1.

3 UNCERTAINTY QUANTIFICATION IN MODELING AND SIMULATION

By definition, a model is an approximation of reality, and the degree of that approximation as well as the various sources of uncertainty affect the accuracy of predictions made with the model. As described in this section, the nature of the problem and the source of uncertainties guide how the UQ is performed.

Generally, uncertainties in modeling and simulation can be associated with

(a) model form
(b) model inputs
(c) numerical solutions
(d) model-basis data

Understanding the differences between these sources of uncertainties is important because different characterization and analysis techniques are often used to quantify the uncertainties originating from different sources and to propagate these uncertainties through the model. Uncertainty propagation is the process of using the knowledge about uncertainty in model input variables to quantify the uncertainty in output variables. In para. 3.1, a brief discussion is provided of the sources of uncertainties in physics-based and empirical models. In para. 3.2, several methods for uncertainty propagation are discussed.

3.1 Sources of Uncertainty in Modeling

There are three sources of uncertainty in physics-based computational models: the model form, the model inputs, and the numerical solution. Uncertainties in the model form arise from assumptions or approximations in the formulation of a specific set of mathematical equations to represent the reality of interest. Model input uncertainties represent the uncertainties in the nondeterministic input variables and parameters associated with the selected model form. Numerical solution uncertainties arise, for example, due to discretization approximations, iterative solution algorithms, and particular computational platform characteristics.

The uncertainties in empirical models originate primarily from three sources: the model form, the model inputs, and the model-basis data. As in the physics-based models, the uncertainties in the model form result from assumptions and approximations in the selection or formulation of a specific set of mathematical equations to represent the reality of interest. The uncertainties in the model-basis data arise from errors, both systematic and random, in the experimental data used to develop the model, as well as from limited sample size and a number of other limitations associated with the model-basis data sets. Unlike in the physics-based models, such uncertainties are a major, and often a dominant, source of uncertainties in the empirical models. Uncertainties associated with the model parameters (and the numerical solution, if applicable) typically play a smaller role in the empirical models than in the physics-based models.

3.1.1 Uncertainties in Model Form. Because the true form of the model is not known, the selection of a specific mathematical form for a given modeling application will lead to some level of model-form uncertainty. In the process of developing a model, there are numerous questions that must be considered, such as

(a) What can and should be modeled mathematically?
(b) What are the important features that the model must accurately represent?
(c) What role do computational constraints, model dimensionality, and code maturity play in the complexity of the model?
(d) What physical principles or data is the model derived from?
(e) Is the model expected to be adequate in the entire prediction or application space?

Model-form uncertainty is generally treated as an epistemic uncertainty because it stems from the inexact nature of the modeling process. Physics intentionally ignored or unintentionally missed in the idealized mathematical representation always exists. One example of missed physics is modeling the boundary condition of a cantilever beam as fully fixed, when in reality there is some rotational movement. Another example of model-form uncertainty involves the computer-aided construction of a model, in which the part geometry is defeatured (simplified by removal of selected details) based on its relevance to the fidelity of the simulation and before meshing. The extent of defeaturing may be driven by engineering intuition, formal geometry feature sensitivity, or both.

In theory, model-form uncertainty can be reduced through model enhancements, such as incorporation of additional mechanisms in a physics-based model or additional terms in an empirical model. An example of adding terms to an empirical model was discussed in para. 2.8, where the base model formulation eq. (2-8) with a single input variable, S, was amended with additional terms representing the effects of additional explanatory variables. The model-form uncertainty contributes to the residual uncertainty in the empirical models. Therefore, reducing the model-form uncertainty by adding appropriate terms to the model formulation results in a reduction of the residual uncertainty, as shown in Figure 2.8-1. Of course, the model-form uncertainty is reduced only if the input variables added to the model formulation have statistically significant effects on the response variable being modeled. It is also recognized that adding fidelity to the model will often introduce additional nondeterministic input variables and parameters, and the uncertainties in these added inputs will contribute to the uncertainty in the model output.

In practice, however, model enhancements have important limitations. Adding new physical mechanisms increases computational expense or may require new code development. Additional physical mechanisms may require modifications to the underlying equations as different mechanisms are often treated with different mathematical formalisms. Many molecular scale mechanisms, for example, are not easily modeled using continuum mechanics. Meanwhile, adding terms to an empirical model may lead to difficulties with simultaneous estimation of many model parameters, commonly referred to as overparameterization or overfitting.

The validation process quantifies the discrepancy between the experimental and simulation outputs. The experimental output with its associated aleatory and epistemic uncertainties is typically taken as the truth (i.e., referent) for the purposes of validation. If all significant uncertainties are accounted for, a sufficient number of samples are used to characterize and propagate the uncertainties, and no mistakes or blunders were made, then the model-form uncertainty will be the main contributor to this discrepancy. Practically, however, the discrepancy between the experimental and simulation outputs always results from multiple sources of uncertainty, only a portion of which is model-form uncertainty. There is currently no widely accepted methodology to isolate the contribution of the model-form uncertainty from the discrepancy between the experimental and simulation outputs.

3.1.2 Uncertainties in Model Inputs. Uncertainties in the model input variables and parameters are typically quantified using statistical methods or expert judgment and available reference information. It is typical to employ a combination of these two methods by having an expert provide an initial assessment of the uncertainty that is then updated using statistical methods when the relevant data become available.

Statistical estimation and model-fitting methods are used for probabilistic UQ where the model input uncertainties are represented by probability distributions. In most cases, the form of the probability distribution (e.g., normal, log-normal) is assumed, and the parameters of the probability distribution are estimated statistically from either available experimental data or lower-level simulations in a hierarchical model. It is noted that empirical models, by definition, involve parameters estimated from the model-basis data, such as b_0 and b_1 in the example discussed in section 2. Unless such parameters are themselves quantities of interest being modeled, they are typically treated deterministically when the model is used to simulate the output variable (see para. 2.5). However, the additional uncertainty in the output variable due to the uncertainties associated with the fitted parameters is incorporated into the residual uncertainty to account for the nondeterministic nature of these parameters (ref. [11]). As a result, the standard deviation of the residual uncertainty is larger than the standard residual error estimated from the model-basis data and is not constant across the ranges of the input variables in the model-basis data set.

Depending on the complexities of the model, quantity, and quality of available data, and the number of input variables and parameters in the model, different statistical methods may be used for the estimation of probability distribution parameters. Experimental data for the uncertain modulus of elasticity, for example, can be used to estimate the mean and standard deviation of its distribution, as illustrated in Figure 3.1.2-1. For situations where the experimental data are scarce, Bayesian inference approaches may be employed. When the model inputs are functions of other variables, lower-tier models are typically required for the uncertainties in such inputs to be properly quantified.

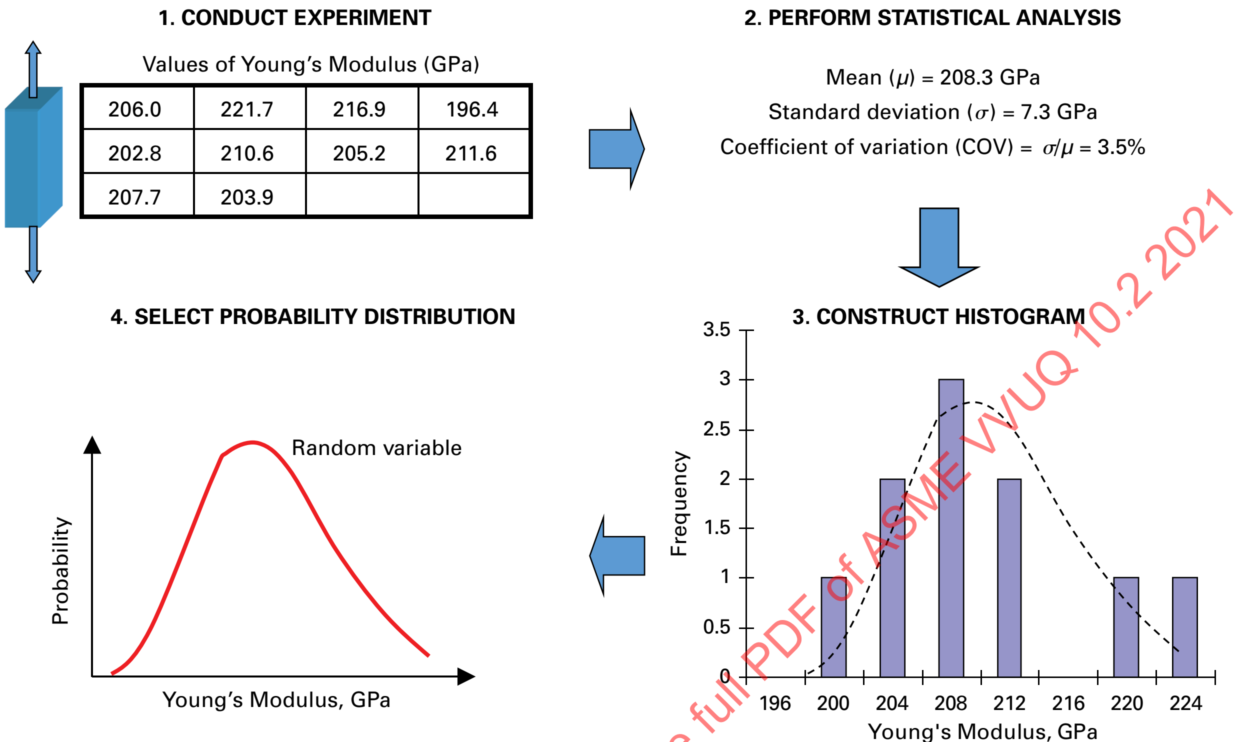
Bayesian estimation (ref. [12]) can be used to update (i.e., improve the estimate of) parameters given new information. The Bayesian methodology has the added benefit of being able to incorporate a wide range of different types of information that other techniques cannot easily consider, such as expert knowledge. As previously stated, uncertainty associated with expert knowledge is epistemic uncertainty. When experimental data are used in Bayesian estimation to improve prior parameter estimates based on expert knowledge, the epistemic uncertainty in the prior estimates is reduced according to the amount of relevant information contained in the experimental data. With sufficient experimental data, the epistemic uncertainty in the Bayesian parameter estimates can be made arbitrarily small; however, there remains aleatory uncertainty in the Bayesian parameter estimates associated with the variability in the experimental data used in the Bayesian estimation (refs. [13], [14]).

When using Bayesian estimation to improve the estimates of model parameters, one must be wary of large changes to the prior estimates relative to the uncertainties originally assumed for those prior estimates. If the changes exceed the range of uncertainties originally assumed, then the changes should be viewed as inconsistent with the prior modeling assumptions, which should therefore be reevaluated. Furthermore, if the updated estimate of any parameter changes significantly from its prior estimate without a correspondingly significant reduction in its uncertainty, that parameter should be removed from the set of parameters being estimated, or additional informative data should be sought.

The uncertainties in the model inputs may also be quantified purely on the basis of expert judgment or reference information from existing standards or available literature. This form of uncertainty quantification is often subjective and may possess a large degree of implied epistemic uncertainty. In such cases, it is advisable to use nonprobabilistic methods, such as interval analysis, so as not to assert knowledge of the probability distribution form and/or its parameters.

3.1.3 Uncertainties in Numerical Solutions. Numerical solution uncertainties arise from a variety of sources including spatial/temporal discretization, solution controls, and computational platform characteristics. Uncertainties generated as part of the numerical solution of the model are termed numerical solution uncertainties and are best addressed by performing calculation (solution) verification. The strategy of addressing numerical solution uncertainties in the overall UQ process is either to demonstrate that they are negligible or to quantify their impact on the model output. When demonstrating that the effect is negligible, the model developer must ensure that this condition is true for the entire space (or range) of parameter uncertainty.

Figure 3.1.2-1
Defining a Random Variable Using Data



Discretization of the idealized geometry for solid mechanics field solvers usually refers to discretization of the defeatured geometry. The resulting level of accuracy in boundary representation of curved entities is usually confounded with the level of spatial resolution and, therefore, spatial convergence of the deformation field. Numerical uncertainty arising from spatial discretization is quantified by refining the mesh while retaining local geometric proportionality between the meshes. Assuming that the RQs are in their asymptotic convergence range with respect to spatial resolution, a Richardson extrapolation process can be used to estimate converged values of the RQs and their convergence rates (refs. [6], [15]).

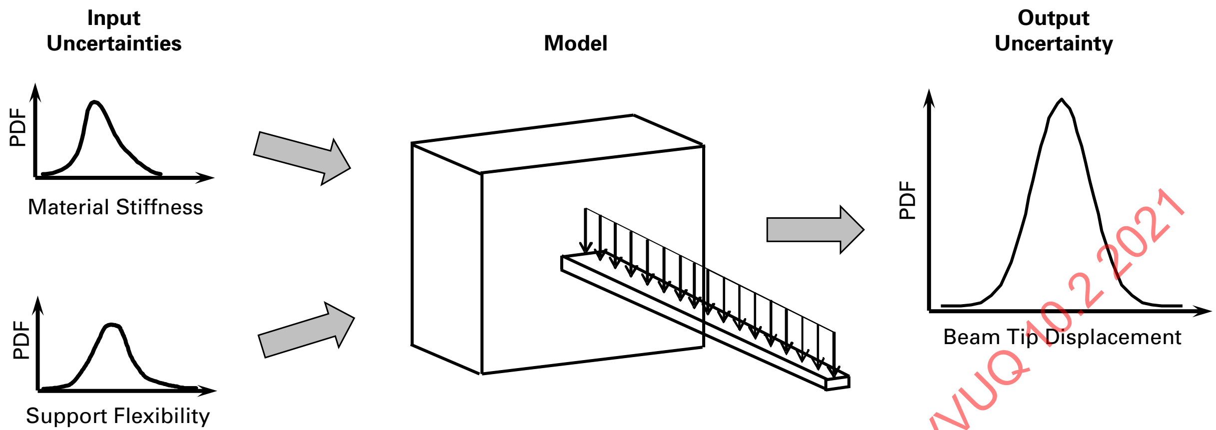
For most load histories and deformation modes of engineering interest, time/load steps are required to resolve the equations of motion as well as path-dependent processes such as incremental plasticity. Additionally, time/load steps may be determined from numerical stability considerations such as the maximum time-step for a given model in explicit dynamics analyses. Typically, solid mechanics modeling systems provide an adaptive time-step size feature that increases or decreases the time-step to improve efficiency. The simplest way of quantifying numerical uncertainties related to temporal convergence is to allow automatic time-stepping in the UQ study and to perform a sensitivity study of the RQs with respect to the maximum allowed time step to demonstrate invariance of the RQs to step size.

Solution controls include, but are not limited to, allowable and maximum residuals for equilibrium iterations, hourglass controls for finite elements, and choices of preconditioners. Hourglass control for a class of problems involving large deformation and failure may be a nonnegligible uncertainty that needs to be included in the overall UQ analysis. As with temporal discretization, the preferred approach is to determine a set of controls for which these factors can be considered negligible.

Numerical uncertainty in spectral analysis arises from the extraction of modes, i.e., the spectral decomposition of system equations of state. The quality of the modes is related to mesh resolution and eigensolver settings, and the number of modes affects the uncertainty in resolving the desired system responses. All these factors may have an effect on the RQs and need to be assessed through a rigorous parametric study.

Computational platform characteristics include microprocessor architecture, the number of processors per node, vendor, operating system, compiler, and compiler optimization levels. In large-scale system simulations on parallel computers, consistency across the various processors must be considered before performing UQ. Invariance of RQs with respect to processor count needs to be evaluated once the computational platform is selected.

Figure 3.2-1
Illustration of the Uncertainty Propagation Process



3.1.4 Uncertainties in Model-Basis Data. Uncertainties in the model-basis data contribute to the overall uncertainty in the model output. Random errors in the model-basis data are one of the sources of the residual uncertainty defined in para. 2.8. The residual uncertainty component associated with these errors is aleatory in nature, since it represents the inherent randomness in the data. With reference to Figure 2.8-1, this component of the residual uncertainty corresponds to the minimum value of the standard residual error that can be attained by incorporating all influential input variables into the model formulation. The systematic errors in the reported values of the output variable may also provide a significant contribution to the uncertainty in the model-basis data, in addition to the contribution made by the random errors and addressed by the residual uncertainty.

In addition to the random and systematic errors in the experimental results, a number of other sources of uncertainty are typically present in the model-basis data. One such source is associated with the substantial lack of balance in the model-basis data with multiple input variables when different experimental conditions are represented by very different numbers of data points. In some cases, the uncertainty arising from this source may be reduced without compromising the model adequacy by means of weighted regression analysis and related statistical techniques. Limited sample size represents another source of uncertainty in the model-basis data, and in engineering practice is typically addressed by increasing the magnitude of the residual uncertainty, e.g., by using the Student's *t*-distribution instead of the standard normal distribution to represent the residual uncertainty (ref. [16]).

3.2 Uncertainty Propagation

Uncertainty propagation, illustrated in Figure 3.2-1, is the process of using knowledge about the uncertainty in the input variables and parameters of a model to quantify the uncertainty in the output variables (i.e., RQs). It is also used to propagate uncertainty from lower tiers in a modeling hierarchy to higher tiers (see section 7). There are numerous approaches to uncertainty propagation. This paragraph briefly discusses the most commonly used and widely accepted approaches.

Uncertainty propagation methods include sampling methods, perturbation methods, and stochastic spectral methods. All of these methods require the assignment of probability distributions for the nondeterministic variables and parameters, typically in the form of PDFs or CDFs. Sampling methods take a variety of forms. The most robust sample-based methods use random sampling from the input variable and parameter distributions and are referred to as Monte Carlo methods (or Monte Carlo simulations). The model is evaluated with each random sample and the results used to construct output PDFs (and/or CDFs). Monte Carlo methods, although robust and accurate, are often computationally expensive, as they require a large number of samples (and thus a large number of model runs) to sufficiently converge on response distributions and statistics. To reduce computational effort, a variety of variance reduction techniques can be used to reduce the number of samples required for convergence. These include stratified sampling, Latin hypercube sampling, importance sampling, and antithetic variates among others (ref. [17]).

Perturbation methods use various numerical integration strategies, typically employing model parameter sensitivities. They were developed for a variety of reasons, but primarily to avoid the computational cost associated with sampling. In many cases, perturbation methods, e.g., first- and second-order reliability methods (ref. [18]), are far more efficient than sampling methods, especially when the model is computationally expensive to evaluate. However, perturbation methods also require considerable computational effort to obtain model parameter sensitivities for problems with a large number of random variables (ref. [19]). There are also hybrid methods that combine sampling and perturbation methods to balance accuracy with computational cost (ref. [20]).

Stochastic spectral (or Galerkin) methods (ref. [21]) such as polynomial chaos augment the system of differential or difference equations that compose the physical model with stochastic parameters that represent the uncertainty. Some of these methods are referred to as intrusive methods because implementation requires intrusive modification of the equations of motion. Other stochastic spectral methods are nonintrusive. While powerful for some applications, they are not discussed here because their use is limited to specific classes of problems and their implementation is complex and requires specialized solvers that are undesirable for practical applications. Sampling methods and perturbation methods are considered nonintrusive because uncertainty is propagated using existing deterministic solvers (e.g., commercial finite element codes) and require no modifications to the governing equations.

All uncertainty propagation methods involve approximations that must be recognized and quantified in the validation process. Moreover, these methods themselves must be verified before being employed in the model validation process.

4 UNCERTAINTY QUANTIFICATION IN VALIDATION EXPERIMENTS

A validation experiment is specifically planned and performed to assess the accuracy of a computational model. Unlike other types of experiments that are performed to improve fundamental understanding of the physical system or to estimate model parameter values, validation experiments are specifically conducted to provide an independent measure to compare with model simulation outputs and quantify the sources of discrepancy. Because experimental outputs are compared with simulation outputs, the ideal validation experiment should be conducted in a highly specified and controlled environment. While this section focuses on validation experiments, the guidance offered here can be applied to other types of experimentation as well.

4.1 Characteristics of Validation Experiments

Validation experiments are planned and performed to generate high-quality experimental data specifically for the purpose of model validation and understanding associated uncertainties (ref. [6]). Full details should be provided in a validation test plan. The applied loads, as-built specimen geometry, materials, initial conditions, boundary conditions, and all other experimental inputs must be precisely controlled or accurately measured. Variables and parameters that are uncertain require a testing strategy that sufficiently represents the input uncertainties, such as replicate testing and design of experiments. Correspondingly, the specimen response must be measured with high, quantified accuracy including uncertainty. A good validation test plan provides as many details as possible, requiring few, if any, assumptions on the part of the model developers' interpretation of the experiment.

The key considerations in planning validation experiments include the following, adapted from ref. [22]:

(a) The experiments should be planned to capture the essential physics of interest, including all relevant physical input variables and parameters as well as initial and boundary conditions.

(b) The experiments should be planned with emphasis on the inherent synergism between the computational and experimental approaches.

(c) The experiments should be planned to enable the uncertainties in the acquired data, including the measurement uncertainty, to be adequately characterized and quantified.

(d) Depending on the problem, multiple RQs should be measured so as to present a range of phenomena and data types (e.g., strain, displacement, acceleration) for comparison with the simulation.

The planning of the experiments should be performed as a joint exercise between the experimentalists and the model developers. This is because the experimentalists need to gain a firm understanding of what the modelers aim to predict, and the modelers need a firm understanding of how the experimentalists intend to measure it. Often, preliminary (i.e., verified but not yet validated) or simplified versions of model simulations can help with the planning of each experiment, but these interactions should be noted in the validation plan and documentation. The final experimental outputs should not be provided to the model developers until after their model simulations have been performed to ensure an unbiased comparison. Alternatively, if the model developers do have access to completed experiments, they should be transparent about any influence on their model simulations that could not be avoided.

There must be a shared understanding of what responses will be measured experimentally. Additionally, there must be an agreement about all relevant physical input variables and parameters as well as initial and boundary conditions to be controlled (and possibly measured) experimentally. It is generally recommended that model developers perform a parametric study with the verified model to determine model sensitivities to help inform the experimental test plan concerning test conditions, instrumentation, data acquisition, and other factors. Pretest sensitivity analyses should also be performed to identify the most effective types of validation experiments, loading locations, sensor locations, and other relevant conditions.

What can be prescribed very precisely numerically may not be readily produced experimentally, and vice versa. For example, the cantilevered boundary condition of Figure 3.2-1 can be prescribed perfectly in the model. However, imposing a true cantilevered boundary condition is impossible to achieve experimentally because there will always be some rotational movement between the test fixture and the test article in the experiment, whereas in the model there will be none. Conversely, using a flat plate to apply a load over a region of a beam may be relatively simple to execute experimentally, but load uniformity is not perfect, which introduces additional challenges for the modeler.

Ideally, the model RQs are directly measured in the validation experiment. However, the experimental RQs may not be directly measurable, may occur in small regions or in regions of high gradients where measurements are not practical, may not be measured without perturbing the intended test conditions, or may not be obtained with a sufficient level of accuracy. Thus, the actual measured RQs may only be related to the desired RQs, and this information must be provided to the modeler prior to performing model simulations.

It is highly beneficial if the experimental data allow different aspects of the model to be assessed. An example is measuring strain at additional points on a beam-bending experiment as opposed to only the maximum deflection. Although some RQs may be of secondary importance, accurate simulations of these responses provide additional evidence that the model correctly simulates the governing physics. This qualitatively builds confidence that the computational model can be used to make accurate predictions for problem specifications that are different (within reason) from those considered in model development and validation.

In many practical cases, experiments involve multiple variables, and validation data under many different testing conditions are required. In such cases, the methodology of statistical experimental design (ref. [23]) often ensures that a minimum amount of experimental work is performed to achieve the required level of statistical confidence in the process of uncertainty characterization and quantification, and that the uncertainty characterization and quantification are performed as efficiently and accurately as possible for a given amount of validation data. Therefore, the methodology of statistical experimental design should always be considered in the planning of validation experiments whenever appropriate and practical.

4.2 Uncertainty Quantification in Validation Experiments

Validation experiments should be performed at multiple validation points to ensure sufficient coverage of the entire validation space. Because uncertainties are involved, a single validation experiment at a given validation point is generally insufficient. Replicate experiments should be performed at each validation point to quantify the uncertainty in the experimental outputs. When replicate experiments are not available, uncertainty in experimental outputs must be estimated by other means, such as relying on experience and judgment. As another example, when symmetry is present, replicate results may be obtained by taking measurements at symmetrical locations on the test specimen.

During the early stages of experimental planning, it is useful to consider all potential sources of uncertainties and to make an estimate of those uncertainties. ASME PTC 19.1 (ref. [24]) recommends performing a pretest UQ followed by a comparison with post-test UQ. Similarly, Coleman and Steele (ref. [25]) reflect this recommendation by performing a general uncertainty analysis early in the experimental planning process, followed by a detailed uncertainty analysis. Estimates are updated as the experimental planning and testing move toward completion and post-test UQ analysis. Many methods can be used for estimating measurement uncertainty, including previous test results, published data, expert judgment, and even comparison of multiple methods of measurement in a single test. Collectively, these pretest and post-test activities promote early and ongoing communication between experimentalists and model developers, identify resource allocation opportunities (both computational and experimental), inform the final stages of experimental planning, likely result in reduced measurement uncertainties, and possibly identify which measurements may dominate in overall result uncertainty.

The result of an experimental test is often calculated indirectly from several direct measurements using either a data reduction equation or a computational simulation. Thus, the uncertainties associated with each individual measurement are propagated to estimate the uncertainty of the result. Techniques for propagation of experimental measurement uncertainty are similar to those used in uncertainty propagation for simulations. For example, Taylor series approximations and Monte Carlo simulation techniques can be used to propagate the measurement uncertainties into the desired form; an in-depth description and multiple examples are given in ASME PTC 19.1.

5 UNCERTAINTY QUANTIFICATION IN MODEL VALIDATION ASSESSMENT

A validation metric is used to compare an uncertain output from the model simulation with a corresponding uncertain output from the validation experiment. This validation metric is then compared to the validation requirement to assess whether or not the validation requirement is satisfied. In selecting the validation metric, key considerations should be what the model is expected to predict (relative to the intended use of the model) as well as what types of experimental data are available or needed.

5.1 Validation Requirements

As defined in V&V 10 (ref. [2]), validation requirements are "specifications and expectations that a computational model must meet to be acceptable for its intended use." They are project-specific and are dependent on budget, schedule, risk tolerance, safety margin, and other considerations. Validation requirements should be agreed upon and specified during the development of the VVUQ plan, prior to performing any validation experiments or model simulations. Validation requirements depend on the intended use of the model, type of analysis, uncertainty quantification approach, available data, and other factors, such as the consequence of making a wrong decision based on a model prediction. For example, a team developing and validating a model to predict automotive crashworthiness may employ an explicit dynamic finite element model to simulate a frontal impact scenario. RQs could include maximum deformation and acceleration histories at various locations on the automobile and the occupant(s). Because validation metrics incorporate uncertainties in both experimental and simulation outputs, validation requirements may need to be defined in terms of uncertainty (e.g., the model is expected to be accurate for a given RQ to within 10% with 90% confidence).

5.2 Validation Metrics and Assessment

Comparison of simulation and experimental outputs, both of which include uncertainties, is performed with one or more RQs. The RQs may be directly computed and measured or may be postprocessed to obtain the desired outputs for comparison. These outputs are labeled "Simulation Outputs" and "Experimental Outputs" in Figure 1.1-1.

Graphical overlays (cross-plots) of simulation and experimental outputs are generally insufficient for purposes of validation assessment. Even the corridor approach, where the model is assumed to be validated if the simulation outputs fall within some specified experimental corridor, is insufficient because the degree of agreement between simulation and experiment is not explicitly quantified.

Comparison of simulation and experimental outputs requires some type of quantitative validation metric, which usually takes the form of a difference measure (ref. [26]). Example metrics may include the difference between the average values of simulated and measured outputs, the difference between statistics of outputs, or even the difference between the probability distribution of outputs. A suitable validation metric should

(a) fully incorporate simulation and experimental uncertainties
(b) quantify the difference between simulation and experimental outputs
(c) reflect the level of uncertainty in the comparison

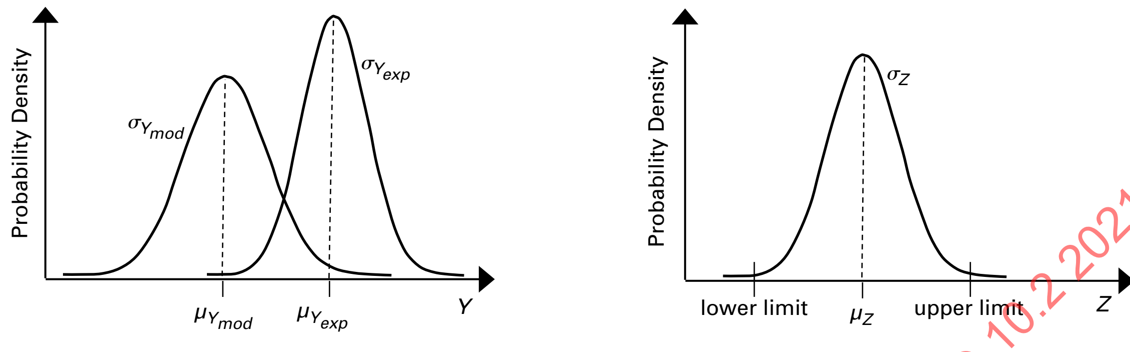
Validation metrics and associated requirements should be established during the requirements definition phase of the conceptual model development and must incorporate both numerical and experimental uncertainties.

Because uncertain quantities are involved, care must be taken to choose the RQs and validation metrics to ensure the model is appropriately and sufficiently challenged. For example, to demonstrate how well the model simulates the measured probability distribution, a metric that quantifies the difference between the simulated CDF and the experimentally measured CDF would be appropriate.

5.2.1 Mean Metric. A simple metric for comparing two uncertain quantities is the difference in the average (or mean) values. An example is presented herein to introduce the concept of a simple metric that considers uncertainty.

Recognizing that the model-simulated RQ is uncertain and characterized as a random variable, Y_{mod}, and that the experimentally measured RQ is also uncertain and characterized as a random variable, Y_{exp}, the mean, μ, for each of these random variables may be computed. The difference between these means, as shown in eq. (5-1), can then be computed and tested against a corresponding validation requirement. This mean, $\Delta\mu$ metric is clearly a function of the random variables Y_{mod} and Y_{exp}, and therefore is a metric that considers uncertainty. However, this single-valued metric does not directly account for (or provide insight into) the magnitude of the uncertainty in Y_{mod} and Y_{exp} and thus little can be stated regarding the uncertainty of this measure.

Figure 5.2.1-1
Extension of the Mean Metric: The Difference Between Y_{mod} and Y_{exp}



(a) Simulated and Measured RQs **(b) Variable *Z***

$$\Delta\mu = \mu_{Y_{\text{mod}}} - \mu_{Y_{\text{exp}}} \tag{5-1}$$

With minimal effort, significant additional information can be obtained by extending the mean metric. If Y_{mod} and Y_{exp} are assumed independent,[2] a new random variable, Z, can be defined as the difference between Y_{mod} and Y_{exp}, from which the mean and standard deviation (σ) of Z can be computed:

$$Z = Y_{\text{mod}} - Y_{\text{exp}} \tag{5-2}$$

$$\mu_Z = \mu_{Y_{\text{mod}}} - \mu_{Y_{\text{exp}}} \tag{5-3}$$

$$\sigma_Z = \sqrt{\sigma^2_{Y_{\text{mod}}} + \sigma^2_{Y_{\text{exp}}}} \tag{5-4}$$

where

$\sigma_{Y_{\text{mod}}}$ = standard deviation of Y_{mod}
$\sigma_{Y_{\text{exp}}}$ = standard deviation of Y_{exp}
σ_Z = standard deviation of Z

This is shown graphically in Figure 5.2.1-1. The simulated and measured RQs are shown in Figure 5.2.1-1, illustration (a), and the new random variable Z is shown in Figure 5.2.1-1, illustration (b).

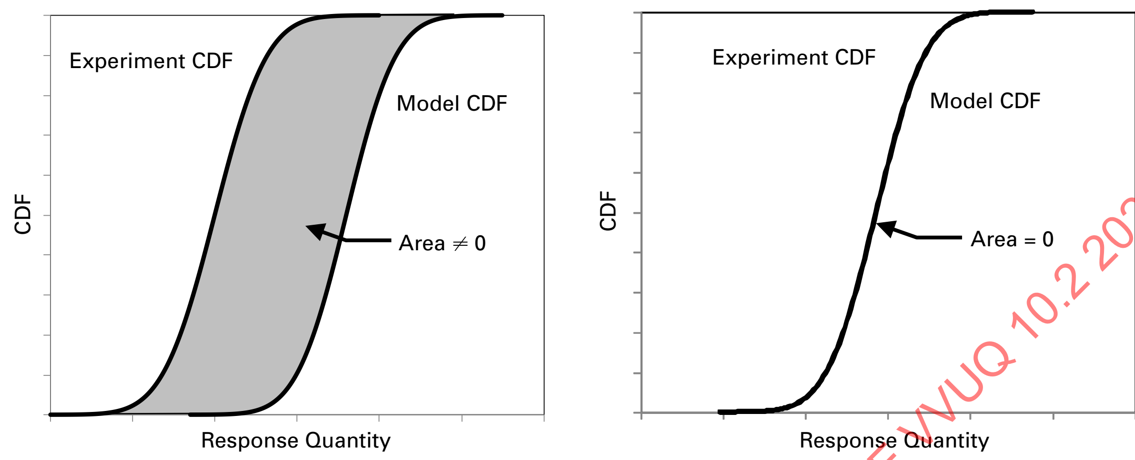
As shown in Figure 5.2.1-1, illustration (b), the variable Z is clearly also uncertain. It is important to note that this metric is different from the mean $\Delta\mu$ metric introduced earlier. As shown in eq. (5-3), the mean of Z is, in fact, equal to the mean metric, but Z also directly reflects the uncertainty in Y_{mod} and Y_{exp}. By quantifying σ_Z, upper and lower limits on Z can be computed at any desired confidence level (e.g., 90%), which can then be tested against a corresponding validation requirement. Therefore, a definitive statement can be made regarding the confidence in which the validation requirement is satisfied.

5.2.2 Area Metric. The area metric possesses many of the desirable features of a validation metric and fully accounts for uncertainties in both the simulation and the experiment (ref. [27]). Accordingly, it provides a good basis for a discussion on the use and interpretation of a metric when uncertainties are involved.

The area metric is a measure of the difference between the CDFs of two random variables, in this case, the CDF of the simulated RQ and the CDF of the experimental RQ. The area metric, Z, is formulated as

[2] Independence between simulated and measured RQs would require that uncertain loads applied to the model and the test article also be independent. This may be accomplished by independently and randomly selecting loads to drive the model and those to drive the test article from the same random source.

Figure 5.2.2-1
Illustration of the Area Metric



(a) Noncoincident CDFs **(b) Coincident CDFs**

$$Z = \int_{-\infty}^{\infty} \left| F_{\mathrm{mod}}(Y) - F_{\mathrm{exp}}(Y) \right| dY \tag{5-5}$$

where

$F_{\mathrm{exp}}(Y)$ = CDF of the experimental RQ
$F_{\mathrm{mod}}(Y)$ = CDF of the simulated RQ

The absolute value in eq. (5-5) denotes that only the magnitude of the difference is of interest, and that all differences between $F_{\mathrm{mod}}(Y)$ and $F_{\mathrm{exp}}(Y)$ should be accumulated in the integration. It can be seen in Figure 5.2.2-1 that the entire distribution is important, not just the mean, standard deviation, or some other summary statistic.

The shaded region in Figure 5.2.2-1, illustration (a), denotes the area computed by eq. (5-5). The two CDFs shown are continuous functions, but the area can be computed even when the experimental or simulated CDF is a stepwise function (individual samples) or deterministic (a vertical line).

An important feature of the area metric is that the area equals zero when the experimental and simulated CDFs are coincident [see Figure 5.2.2-1, illustration (b)]. Any difference between the two CDFs results in some area being accumulated in eq. (5-5). This is appealing and makes intuitive sense from the standpoint of a metric being a distance measure; when two points (or functions in this case) are coincident, the distance between the two points must equal zero. However, when uncertainties are involved, care must be taken to ensure that this result is interpreted correctly. An area of zero means that the model is simulating the same uncertainty as was measured in the validation experiment. From one point of view, the model cannot be expected to do any better than this. However, from the perspective of the ultimate intended uses of the model, further model development may be required to reduce the uncertainty associated with the model prediction.

5.2.3 Error Metric. The error metric is defined as the relative difference between the uncertain simulation output, Y_{mod}, and the uncertain experimental output, Y_{exp}, written as

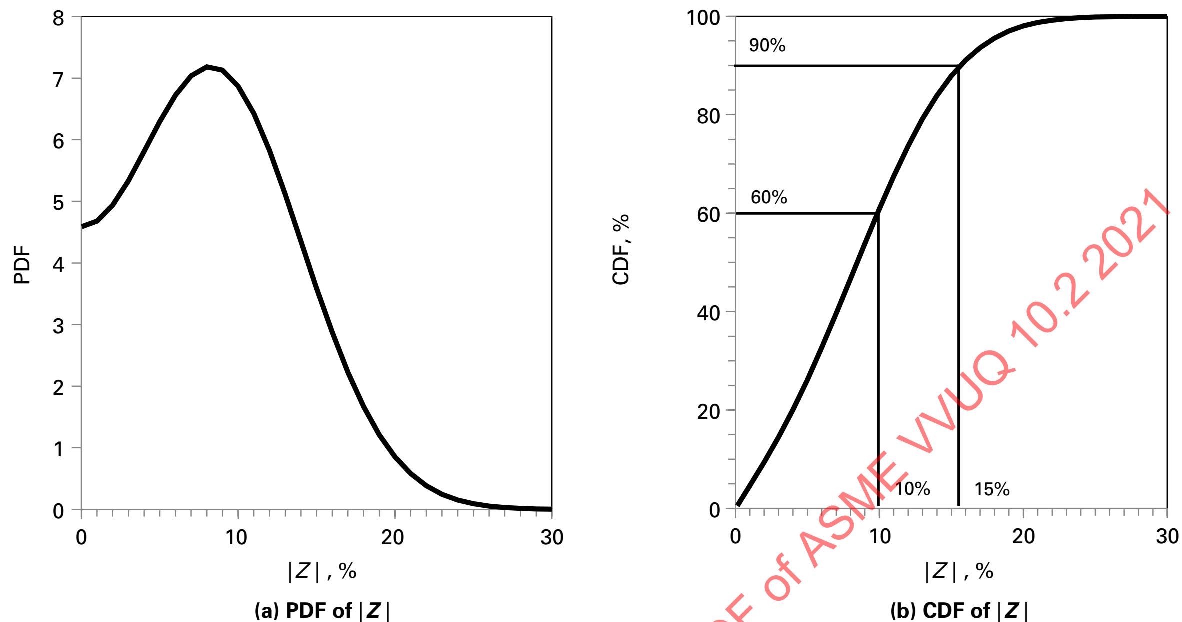
$$Z = \frac{Y_{\mathrm{mod}} - Y_{\mathrm{exp}}}{Y_{\mathrm{exp}}} \tag{5-6}$$

where Z is valid when the realizations of Y_{exp} are not zero (ref. [28]).

Because Y_{mod} and Y_{exp} are both random variables, Z will also be a random variable (i.e., uncertain). The PDF of Z [as in Figure 5.2.3-1, illustration (a)] will usually not follow any standard PDF and will, in general, need to be evaluated by sampling or some other uncertainty analysis method.

The probability of the absolute value of the error being less than or equal to a particular error value, z, is given by

Figure 5.2.3-1
PDF and CDF of the Error Metric



$$p = P(|Z| \leq z) \tag{5-7}$$

which by definition is the CDF of $|Z|$ [as in Figure 5.2.3-1, illustration (b)]. The presence of the absolute value in eq. (5-7) reflects that the sign of the error is typically unimportant in the model validation problem; our interest is focused only on the magnitude of the difference between the possible values of Y_{mod} and Y_{exp}.

It should be noted that Y_{mod} in eq. (5-6) describes the set of possible model simulations of a particular RQ, and in the typical case is the result of propagating input uncertainties (random variables) through the model. Likewise, Y_{exp} describes the set of possible outputs from the validation experiment, which is typically a set of replicate experiments, since one or more uncertain test inputs are involved (see para. 4.2). The CDF for Y_{exp} may also represent propagation of uncertainties from individual measurements that make up the overall experimental result.

If the CDF of Y_{mod} overlays the CDF of Y_{exp} exactly (i.e., they are coincident), then clearly Z given in eq. (5-6) will still be a random variable, and there will be many combinations of p and z that satisfy eq. (5-7). This is because p represents the sum of all possible errors for a given z, which is equivalent to drawing a random sample from Y_{mod} and Y_{exp}, computing and accumulating the error, and repeating these steps a large number of times until convergence. It can be shown that the error is a minimum—but not zero—when the CDF of Y_{mod} and Y_{exp} are coincident. For Z to equal zero requires that the standard deviation of Y_{mod} and Y_{exp} also be zero, i.e., deterministic.

The PDF and CDF of $|Z|$ are illustrated in Figure 5.2.3-1. The shapes of the PDF and CDF generally follow what is called a folded distribution, which is the result of the absolute value in eq. (5-7).

The CDF can be used to return a probability given an allowable error, or vice versa. Figure 5.2.3-1, illustration (b), shows that a 10% error corresponds to a 60% probability. Thus, there is a 60% probability that the error between model and test is not greater than 10%. If a higher probability is desired, say 90%, then the corresponding error is 15%. If the validation requirement is 10%, then the model is not validated. However, if a probability of 60% or less is deemed acceptable, then the validation requirement is met.

The mean metric quantifies the difference in the mean of the simulated and measured RQs, the area metric quantifies the difference between the simulated uncertainty and the measured uncertainty, and the error metric quantifies the percent error between the uncertain simulation output and the uncertain experimental output. The mean and area metrics equal zero when the simulated and experimental CDFs are coincident, whereas the error metric equals zero only when the CDFs are coincident and the uncertainty in the simulated and measured RQs is zero. The differences