

The JANNAF Simulation Credibility Guide on Verification, Uncertainty Propagation and Quantification, and Validation (Invited)

Unmeel Mehta^{*}

NASA Ames Research Center, Moffett Field, California

Vicente Romero[†]

Sandia National Laboratories, Albuquerque, New Mexico

Dean Eklund[‡]

United States Air Force Research Laboratory, Wright-Patterson Air Force Base, Ohio

Jeffrey Pearce[§]

Universal Technology Corporation, Dayton, Ohio

and

Nicholas Keim^{**}

The Johns Hopkins University, Columbia, Maryland

To make simulation-based critical decisions, decision makers need to know the level of credibility of the simulation for the intended use. Credibility depends on the accuracy of simulation, which is quantified by uncertainty. Verification and validation are processes that assess simulation accuracy. Simulation verification, propagation and quantification of uncertainties, and referent data or circumstantial evidence, are essential to establishing simulation credibility. Under the auspices of the Joint Army-Navy-NASA-Air Force (JANNAF) Interagency Propulsion Committee, a guide for assessing simulation credibility is being developed. Simulation verification, uncertainty propagation and quantification, and simulation validation are discussed. Some state-of-the-art approaches are explained and demonstrated. These approaches are anticipated to be useful to other technical communities conducting physics-based simulations.

I. Introduction

*“The mere formulation of a problem is often far
more essential than its solution . . .”*

— Albert Einstein

In 1953, Kawaguti obtained a numerical solution of the Navier-Stokes equations for the flow around a circular cylinder at Reynolds number 40 using a mechanical desk calculator, working 20 hours a week for 18 months.¹ During the last sixty years, tremendous advances in many key facets of digital computing from hardware to algorithms have taken place. These advances are already making a significant impact on aspects of design, engineering, manufacturing, and safe operations of aerospace components and sub-systems. In other endeavors, such as nuclear power generation, nuclear weapon safeguards, civil

^{*}Division Aerospace Engineer, NASA Supercomputing Division, MS 258-5, AIAA Associate Fellow.

[†]Model V&V, UQ, and Simulation Credibility Dept., Sandia National Laboratories, AIAA Senior Member.

[‡]Senior Aerospace Engineer, High Speed Systems Division, AIAA Senior Member.

[§]Program Manager, Aerospace Systems Division.

^{**}Associate Research Engineer, Chemical Propulsion Information Analysis Center, AIAA Senior Member.

engineering, healthcare, and science, such computational advances are also making a tremendous impact. These advances have the potential to revolutionize the way critical decisions are made and to be a key element to advance science and technology—if the uncertainty of a simulation is quantified and the amount of uncertainty is determined to be acceptable for the intended use. Simulation credibility has been the principal issue from the inception of numerical simulation technology.²

Simulation credibility depends on the accuracy of simulations for the intended use. How accurate are simulations? What level of confidence can one assign to simulations? Is the available evidence enough to make correct decisions? Such questions are the critical ones for which decision-makers need answers. If simulations are presented for making decisions, then there has to be a statement regarding the credibility of simulations. It is essential for risk-based decision-making and risk-informed decision analysis. The burden of proof for simulation credibility lies with the presenter of simulations.

The Modeling and Simulation Subcommittee (MSS) of the Joint Army-Navy-NASA-Air Force Interagency Propulsion Committee was established in 1999. Under the auspices of MSS, a voluntary effort began in 2007 to develop a simulation credibility guide to describe and demonstrate some state-of-the-art approaches for simulation verification, uncertainty propagation and quantification, and simulation validation. Seven Simulation Credibility Guide Workshops were held primarily to identify and discuss advanced simulation credibility approaches and, in part, to highlight key modeling and simulation issues related to air-breathing hypersonic propulsion, and liquid and solid rocket propulsion.^{3,4,5}

Herein, the Guide effort is summarized. The philosophy of simulation credibility assessment is presented, selected advanced approaches are summarized for simulation verification, uncertainty propagation and quantification, and validation, and these approaches are demonstrated with examples. The presented credibility philosophy and methods are anticipated to be useful to other technical communities conducting physics-based simulations.

II. Background

“The beginning of wisdom is to call things by their right names.”

– Chinese proverb

A. Simulation Credibility

In making critical decisions based on simulations, the estimated simulation error (including model error, numerical error, and input parameter error) is principally relevant. This estimate of error is essential for establishing simulation credibility. The acceptability of this estimate for the intended use leads to the acceptability of the simulation model. The simulation model could include a physics model, calibrated physics model, or a surrogate (meta) model.⁶ Model calibration is the act of tuning a simulation model, somewhat without scientific justification.⁷ A meta model is a model of a model. Irrespective of the kind of model used, to determine simulation credibility a comparison of simulation is done with a referent. Simulation validation is the process to determine the level of accuracy or credibility of the simulation.

The AIAA and ASME validation guidance is that there can be no validation without experimental (real-world) data with which to compare the simulation or the derived simulated result.^{8,9,10} However, microphysics valid simulations with estimated error in macrophysics quantities of interest may be used as a referent to validate macrophysics simulations. Sub-grid-scale models for Large Eddy Simulations (LES) can be developed and validated with Direct Navier-Stokes Simulations (DNS). Likewise, simulations generated with a physics model serves as a referent to validate simulations based on a meta model.⁶

The validation of a real-world model must be done with all necessary quality real-world data (referent data, experimental/test data). Nevertheless, frameworks, techniques, processes, etc. used for validation are valuable and are appropriate for validating any simulation with credible referent data.

Herein, the focus is on simulation credibility. *Validation assesses simulation accuracy, with respect to a referent with acceptable estimated error, from the perspective of intended uses of the simulation.* A favorable outcome of validation may lead to the use of the validated simulation model for predictions. However, a valid mathematical model does not necessarily assure a credible or acceptable simulation.

Simulation validation is conducted at critical and sensitive locations in the green domain (Fig. 1) identified with Design of Experiments (DoE).⁶ If the outcome of simulation validation is satisfactory,

predictions are done at locations other than these validation locations. The credibility of predicted simulations inside the model validation domain and in the neighboring region (blue domain in Fig. 1) could be higher than outside these domains. To differentiate predictions in the model validation domain and those outside this domain, Ref. 6 introduced the word “postdiction” for predictions in the model validation domain and uses the word “prediction” elsewhere.

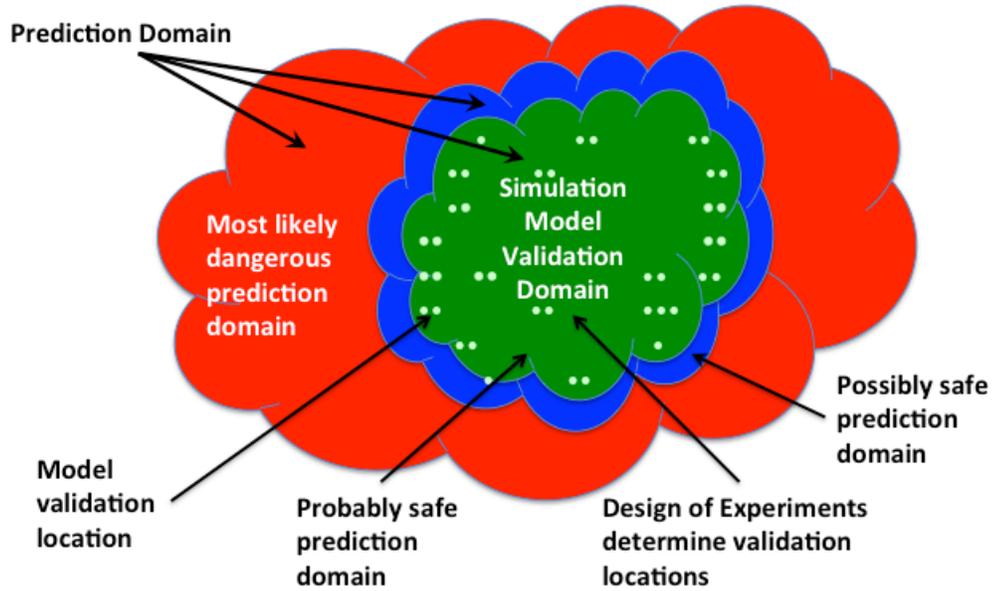


Figure 1. Simulation validation domain.

Figure 2 presents a high-level diagram of the important role of uncertainty in various aspects of simulation validation. The top row and the center of the diagram show a linkage between the requirement for simulation, resource allocation, and simulation validation for simulation model validation and prediction validation. These aspects play major roles in the credible use of predictions for design, analysis, and operations of a system and its components.

A proliferation of research has been conducted in advancing the state of the art of forward propagation of input/parameter uncertainty and quantifying the output uncertainty. The assessment of model and parameter uncertainties, that require the inverse assessment of uncertainties, is often a much more difficult problem than forward uncertainty propagation and subsequent quantification.

The International Organization for Standardization (ISO) has codified the concepts and definitions of error and uncertainty for the experimental community in the Guide to the Expression of Uncertainty in Measurement (GUM).¹¹ Coleman and Stern adapted most of these concepts and definitions for simulations.¹² Subsequently, ASME V&V 20 did the same. A common language between the experimental community and the simulation community is preferred.

During simulation validation, a comparison of simulation is done with a referent. The comparison error equation adapted from ASME V&V 20 is as follows:¹⁰

$$E = S - R = (T + \delta_{Sval}) - (T + \delta_R) = \delta_{Sval} - \delta_R \quad (1)$$

Here, E , S , R , and T are the comparison error, simulation, referent, and true value, respectively. The latter value is assumed to be the same for simulation and for referent. The quantities δ_{Sval} and δ_R are, respectively, simulation error and referent error during validation.^{††} The simulation error, δ_S , is

^{††} A more general comparison scheme that handles added complications of multiple replicate experiments with randomly varying stochastic systems and models for predicting the stochastic phenomena is discussed in Section V.

$$\delta_S = \delta_{mod} + \delta_{num} + \delta_{input} \quad (2)$$

where quantities δ_{mod} , δ_{num} , and δ_{input} are, respectively, model error, numerical error, and input parameter error and $\delta_{num} = \delta_{numm} + \delta_{numd}$ with δ_{numm} and δ_{numd} , respectively, numerical method error and discretization error. The model error is:

$$(\delta_{mod})_{val} = [E - (\delta_{num} + \delta_{input} - \delta_R)]_{val} \quad (3)$$

A measure of uncertainty is defined as an estimate of an interval $\pm u$ that should contain error, δ . The model error falls in the interval given by

$$(\delta_{mod})_{val} \in [E - u_{mod}, E + u_{mod}]_{val} \quad (4a)$$

where E is an estimate of δ_{mod} and with:

$$(u_{mod})_{val} = \sqrt{(u_{num}^2 + u_{input}^2 + u_R^2)_{val}} \quad (4b)$$

The model uncertainty, u_{mod} , is a systematic standard uncertainty and is defined as an estimate of the standard deviation of the parent population from which δ_{mod} is a single realization. No assumption about the form of the parent distribution is associated with the definition of u_{mod} . It is determined from a combination of uncertainties in numeric, input, and referent. Simulation verification provides u_{numd} and uncertainty propagation of input uncertainties provides u_{input} . Thus, simulation verification and input uncertainty propagation are necessary first steps for establishing simulation validity. The referent uncertainty, u_R , must also be estimated. Please note that $(\delta_{mod})_{val}$ is generally not a constant for different uses of the simulation model.

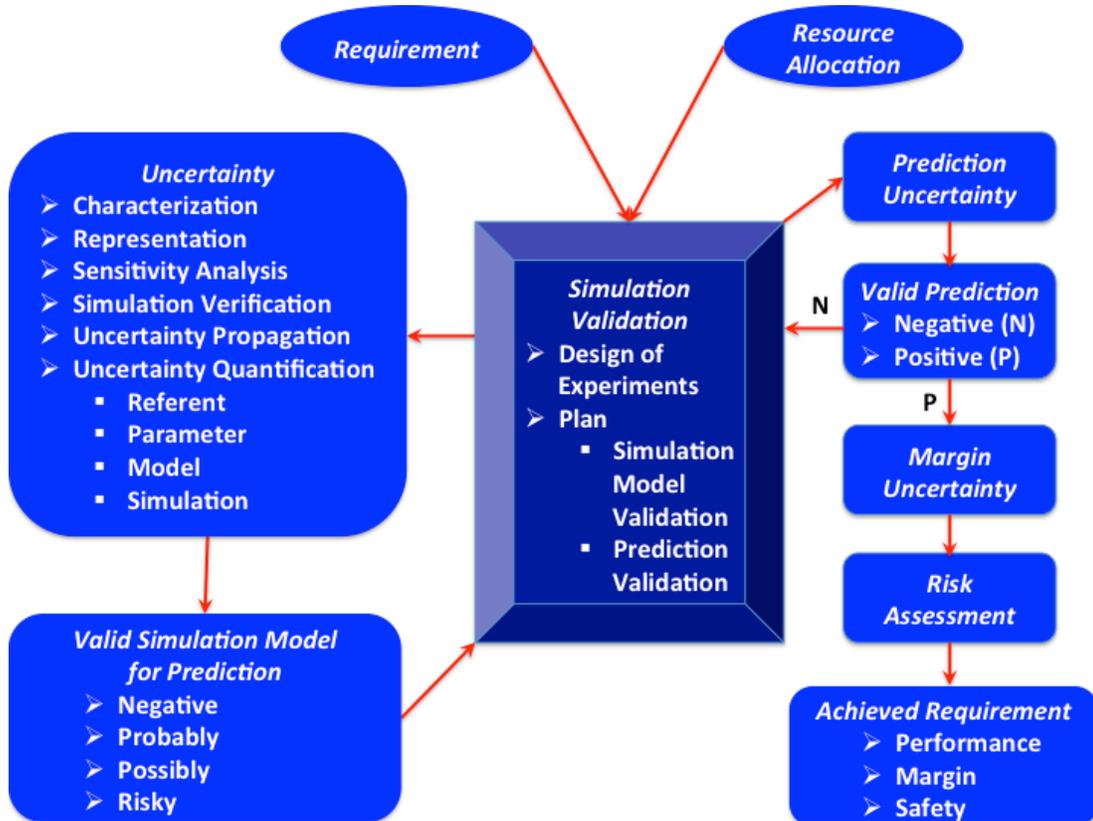


Figure 2. The role of uncertainty in simulation validation for a simulation-based design.

A mathematical model (original model), and the corresponding simulated mathematical model (modified model) are not the same.¹³ When the mathematical model is difficult or impossible to solve analytically, a modified model is solved non-analytically, for example, using computational fluid dynamic techniques. If the mathematical model is $f(u) = 0$, then the modified model is $f(u) + g(u,h) = 0$, where h represents some measure of the discretized domain (such as characteristic cell size). A great deal of effort is spent on reducing the effect of $g(u,h)$. This effect cannot be eliminated. Simulations are generated with a model that is not a mathematical real-world model.

Simulation model validation has other issues. Multi-dimensional problems are often difficult to solve, and simulations could be non-deterministic. There is the possibility of achieving an erred simulation during mesh refinement for simulation verification (e.g., Ref. 14). The mathematical model may include multiple physics models. An example of multiple physics models in a single simulation problem is provided in Fig. 3. Not knowing all necessary input conditions, having only partial rather than all relevant real-world data, or conducting validations at a very few locations in the model validation domain further complicates the validation effort. Different test facilities and measurement techniques may provide somewhat different data. The assumed true value of measurement determines the estimate for measurement error (Fig. 4). Operational environment of a system may differ from the environment in a test facility. Finally, limited resources could affect whether or not the simulation validation is conducted properly.

The error in simulation model, δ_{mod} , is an aggregate of all modeling errors, including those from different physics models and from non-physical (numerical) models that interact or influence the physics models. It is often difficult to completely disaggregate these modeling errors.

Although the NS equations model is valid for a class of real-world physics, DNS require a referent to establish the credibility of such simulations. Strictly speaking, these simulations are with modified NS equations. A confirmation through simulation verification requires that numerical errors, both spatial and temporal, are negligible and all pertinent flow scales (in frequency domain) up to viscous dissipation are resolved. Physics with multiple time and space scales are difficult to simulate given the limitation of computing technologies and tools. These simulations are dependent on inputs such as initial and boundary conditions that are often not known precisely.

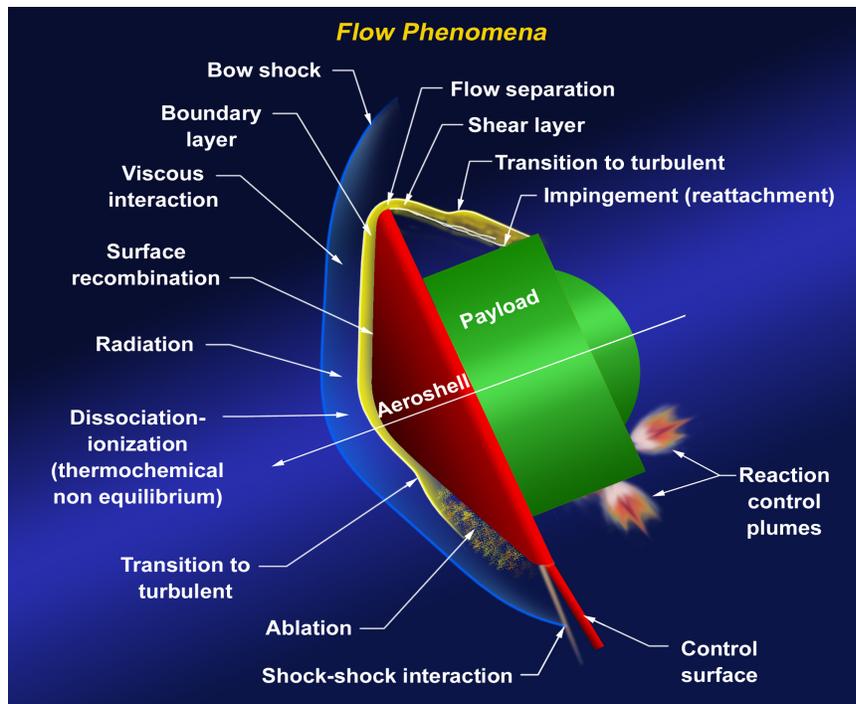


Figure 3. Kaleidoscope of physics surrounding a craft during atmospheric entry.

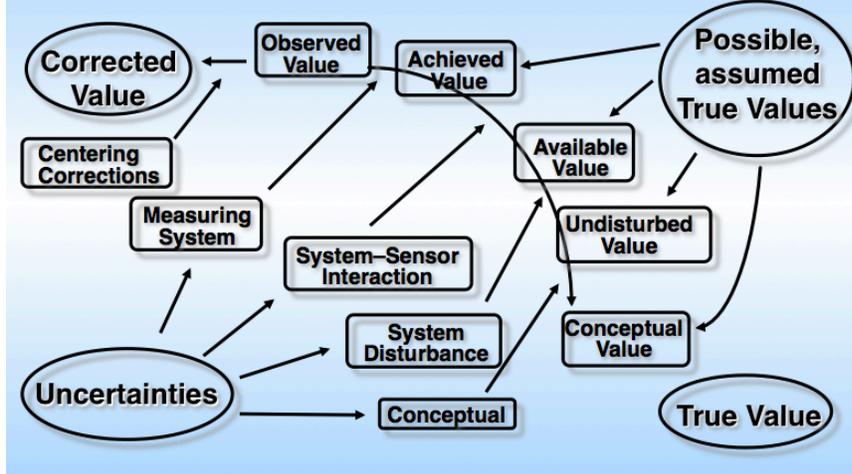


Figure 4. Observed, true values & uncertainties in tests (adapted from Ref. 15).

The user of simulations is interested in the estimated simulation error. When model validation is conducted, the simulation error is:

$$\delta_S = E + \delta_R \quad (5)$$

The validation simulation error, δ_{Sval} , falls in the following interval

$$\delta_{Sval} \in [E - u_{val}, E + u_{val}] \quad (6a)$$

where:

$$u_{val} = \sqrt{u_R^2} \quad (6b)$$

If the simulation error [Eq. (6a)] is acceptable, then the simulation is valid for the intended use as defined for validation. Additionally, if the outcomes of simulation verification and uncertainty propagation are acceptable, then the model error [Eq. (4)] is acceptable. Then the simulation model is valid. If the simulation model is valid, then the modified mathematical model is valid for the numerical method used and input parameters considered.

Model developers are principally interested in reducing δ_{mod} . The decision maker or the user of simulations is principally interested in predictions. This person needs to know simulation prediction error, δ_{Spred} , for making critical decisions. Prediction validation is conducted with circumstantial evidence.^{6,7} Conceptually, prediction error is the interval given by Eq. (7). This equation is used only when the underlying statistical assumptions regarding independent uncertainties are valid.

$$\delta_{Spred} \in [E_{pred} - u_{pred}, E + u_{pred}] \quad (7a)$$

with:

$$u_{pred} = \sqrt{(u_{mod}^2 + u_{input}^2 + u_{num}^2)_{pred}} \quad (7b)$$

In Eq. (7b), u_{pred} is the root sum square of the standard model, numeric, and input uncertainties.

The prediction uncertainty may be determined in four ways with varying degree of circumstantial evidence. If the prediction uncertainty is acceptable for intended use of the prediction, then the prediction is acceptable. In any case, the simulation code and its use ought to be the same for predictions as for model validation in order to make use of these model validation results during prediction. The following are the alternatives:

1. u_{pred} is assumed to be the same as u_{val} .
2. u_{num} and u_{input} are computed and E and u_{mod} are
 - a. assumed to be the same as those at the nearby validation location,
 - b. interpolated to the prediction location within the model validation domain, or
 - c. extrapolated to the prediction location outside the model validation domain.
3. E , u_{mod} , u_{num} and u_{input} are interpolated to the prediction location within the model validation domain.
4. E , u_{mod} , u_{num} and u_{input} are extrapolated to the prediction location outside the model validation domain.

Whichever alternative is selected, it must be justified. Interpolation or extrapolation is done from neighboring validation locations. In the second approach, numerical and input errors, δ_{num} and δ_{input} , are determined. Generally, the third approach is much less risky than the fourth approach. The extrapolation error can be dominant. This error is very difficult to quantify.

B. Some Simulation Credibility Frameworks

The significance of processes for establishing simulation credibility is manifested in the creation of guides, standards, archival publication journals, and professional society conferences. Table 1 presents some existing definitions of verification and validation.

Table 1. Some Existing Definitions of Verification and Validation.

Source	Verification	Validation
AIAA-G-077-1998 ⁸	The process of determining that a model implementation accurately represents the developer's conceptual description of the model and the solution to the model.	The process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model.
Mehta ⁶	Verification assesses the credibility of a simulation model by estimating the degree to which this model is an accurate representation of the conceptual model from the perspective of the intended uses of simulations.	Validation assesses the credibility of a reality model or a simulation by estimating the degree to which it is an accurate representation of reality from the perspective of intended uses of the model or simulation.
Romero ¹⁶		The compilation of useful indicators regarding the accuracy and adequacy of a model's predictive capability for output quantities (possibly filtered and transformed) that are important to predict for an identified purpose, where meaningful comparisons of experiment and simulation results are conducted at points in the modeling space that present significant prediction tests for the model use purpose.
DoDI 5000.61 ¹⁷	The process of determining that a model or simulation implementation and its associated data accurately represent the developer's conceptual description and specifications.	The process of determining the degree to which a model or simulation and its associated data are an accurate representation of the real world from the perspective of the intended uses of the model.

The definitions used for verification and validation in ASME V&V 10 and V&V 20 are consistent with those in AIAA-G-077-1998.⁸⁻¹⁰ The definition of verification used in NASA-STD-7009¹⁸ is consistent with those used in AIAA and ASME documents; whereas, the definition of validation used in this NASA standard is consistent with that of Mehta.⁶

AIAA, ASME, and DoDI define the word “verification” without the following qualification: “for the intended use.”^{8-10,17} This phrase is necessary because the required level of numerical accuracy depends on the intended use. For example, the level of numerical accuracy and associated uncertainty required for the drag coefficient of a commercial transport aircraft is quite stringent compared to that required for its lift coefficient for the same modeling accuracy. Mehta and NASA-STD-7009 use this phrase in their definition of verification.^{6,18}

The AIAA, ASME, and NASA documents consider both code and simulation verification under the definition of “verification.” But code verification and simulation verification are addressed separately. In DoDI, verification only verifies whether the code is built correctly. Mehta⁶ considers only simulation verification; code verification is a very important and necessary but a separate task.

AIAA-G-077-1998, ASME V&V 10, ASME V&V 20, and DoDI do not use the word “simulation” in the definition of validation. Mehta,⁶ NASA-STD-7009, and DoDI include the word “simulation” in their definition of validation: a model is valid, only if the simulation resulting from the model is valid. A necessary first step is simulation validation to validate a model. In DoDI, validation assesses both numerical and modeling accuracy.

Romero considers implementation issues in model validation within a connected context of experiments to extrapolative predictions and provides an “operational definition” of model validation (Table 1).¹⁶ The metric of this definition differs significantly from those of other definitions of validation in Table 1. The “Real Space” model validation metric and method to execute this operational definition of validation are discussed and illustrated in Section V.

NASA-STD-7009 presents a Credibility Assessment Scale and states: “the credibility assessment scale does not purport to measure credibility; rather, it assesses the [models and simulations (M&S)] results, and the rigor of the processes used to produce them, against key factors that affect the credibility judgment.”¹⁸ The burden of assessing simulation credibility is left to the decision maker. This person is made responsible for this assessment in addition to his responsibility to make the decision based on the credibility of simulation.

This standard requires that “reported uncertainty estimates shall include one of the following: (1) a quantitative estimate of the uncertainty in the M&S results; or (2) a qualitative estimate of the uncertainty in the M&S results; or (3) a clear statement that no quantitative or qualitative estimate of uncertainty is available.”¹⁸ A statement on the uncertainty in the results is required. But this standard does not mandate the quantification of uncertainty, which is necessary to estimate the level of accuracy.

Eight factors are evaluated. These are Verification, Validation, Input Pedigree, Results Uncertainty, Results Robustness, Use History, M&S Management, and People Qualifications. Each factor is evaluated with a five-level defined assessment scale. How to conduct simulation verification, simulation validation, and uncertainty propagation and quantification are not explained nor are related best practices identified.

The NASA policy is to “prefer use of performance (outcome-based) standards in procurement activities over design or process (method-based) standards.”¹⁹ But the focus of NASA-STD-7009 is on documentation, output, and processes rather than on outcome. It lists 49 requirements of which 36 requirements assert, “shall document,” and three requirements instruct to “report.”¹⁸

Two key shortcomings of the NASA standard are the credibility assessment approach (logic and acceptability of non-uncertainty-based assessment) and the excessive burden of documentation. The application of this standard does not lead to a credible simulation.

The fourth generation Predictive Capability Maturity Model (PCMM) is a comprehensive framework for managing the risk in the use of models and simulations.²⁰ PCMM is “an expert elicitation tool designed to characterize and communicate completeness of the approaches used for computational model definition, verification, validation, and uncertainty quantification associated for an intended application.” The principal objectives of a PCMM evaluation are to accurately and transparently communicate computational simulation capability, and to develop input for effective planning. PCMM assists in avoiding the following types of mistakes in models and simulations: false negative or false positive assessment of the model, solving the wrong problem, and incorrect use of M&S information. The PCMM approach could be used to determine the validity of a simulation model at prediction location with circumstantial evidence.

The following nine factors are assessed in the PCMM:

- Customer Specification Completeness
- Code Verification
- Representation and Geometric Fidelity

- Solution Verification
- Validation – Hierarchy
- Experimental Data for Constitutive Model Calibration
- Experimental Data for Computational Simulation Model Validation
- Validation – Component
- Uncertainty Quantification

Each factor has sub-factor descriptors and each descriptor consists of four levels of evidence requirement. The focus of PCMM is on evidence-based assessment and on evaluation of evidence quality.

The PCMM framework is recommended for all physics-based modeling and simulation efforts. However, presently, PCMM does not address two issues: regulatory requirements and numerical methods.²⁰ The regulatory requirements may dictate a more systematic and definitive assessment of simulation credibility than is feasible with PCMM. Numerical methods may impact simulated physics and modify the effect of or interact with physics model(s) used in simulation. For example, methods of uncertainty propagation may impact the physics simulated with the simulation model.

Difficulties in properly conducting verification and validation, the inability to properly quantify uncertainties, the lack of real-world data, the limitations of computing technologies, and the limitations of resources have led to the consideration of circumstantial evidence and, in part, to the development of non-deterministic perspectives and methods, first introduced in Ref. 21. Ultimately, credible predictions are needed. Circumstantial evidence plays a major role in establishing this credibility.

Non-deterministic approaches are used to manage simulation uncertainty and to provide decision-support information for addressing high-consequence failure events associated with the development and operation of systems. Both probabilistic and non-probabilistic methods are used to propagate, quantify, and manage uncertainty. Often values of uncertainties are assumed because they are difficult to quantify. The credibility of a non-deterministic simulation for making a critical decision is a primary concern.

III. Simulation Verification

“The man of science has learned to believe in justification, not by faith, but by verification.”

— Thomas Henry Huxley

The Simulation Credibility Guide presents three approaches for conducting simulation verification. The focus for simulation verification is on estimating discretization errors, and in particular, spatial discretization error estimation. Rider, et al. survey the state-of-the-art of simulation verification.²² They illustrate the application of Richardson Extrapolation (RE) to quantify uncertainty in integral output quantities. Cavallo advances Zhang’s Error Transport Equations (ETE) method.^{23,24} This method is used not only for quantifying discretization error but also for adaptive meshing. Nemeč and Aftosmis demonstrate the power of adaptive mesh refinement with adjoint-based error estimates.²⁵ They present results showing reliable estimates and automatic control of discretization error at an affordable computational cost.

A. Simulation Verification Roadmap and Richardson Extrapolation

Rider, et al. identify four general types of simulation verification errors: (1) round-off errors, (2) sampling errors, (3) iterative solver errors, and (4) discretization errors. Round-off errors are an unavoidable consequence of representing numbers on computers with finite precision, and are typically a small contribution to numerical error. Sampling errors are present for stochastic methods (e.g., Monte Carlo methods). Iterative solver errors are a consequence of solving any linear or non-linear system of equations during the course of the computational procedure. Discretization errors are associated with the numerical scheme used to obtain a simulation to the discrete equations that are an approximation to the mathematical equations on a discrete domain. In this work, truncation errors are synonymous with discretization errors.

Rider, et al. have proposed a workflow for guiding simulation verification analyses:

1. Starting with an algorithm implementation (i.e., code) that has passed the appropriate level of software quality assurance and code verification, choose the software to be examined.
2. Provide an analysis of the implemented numerical method including accuracy and stability properties from the code verification analysis.
3. Produce the code input to model the problem(s) of interest.
4. Select the sequence of mesh discretizations to be examined for each problem, and the input necessary to accomplish these calculations.
5. Run the code and provide the means of producing appropriate metrics to evaluate the difference between the computed solutions based on numerical parameters within the control of the code user. This can also include the numerical method chosen (order of approximation).
6. Use the comparison to determine the sequence of estimated errors corresponding to the various discretizations and tolerances.
7. The error sequence allows the determination of the rate-of-convergence for the method, which is compared to the theoretical rate. For iterative solver errors, the error is a function of the stopping criteria and the discretization.
8. Using these results, render an assessment of the accuracy (level of error estimated) for the simulation for a given set of numerical settings.
9. Verify the degree of coverage of features in an implementation by testing.

Consider a characteristic grid spacing convergence study where 3, 2, and 1 represent the coarse, medium and fine grids with spacing Δx_3 , Δx_2 , and Δx_1 .^{††} Now let

$$r = \frac{\Delta x_2}{\Delta x_1} = \frac{\Delta x_3}{\Delta x_2} \quad (8)$$

where r in the following is assumed to be constant, but is not required. Also let S_3 , S_2 , and S_1 represent the solutions on the coarse, medium and fine grids. The solution changes and convergence ratio, D , are:

$$\Delta_{32} = S_3 - S_2 \quad (9)$$

$$\Delta_{21} = S_2 - S_1 \quad (10)$$

$$D = \Delta_{21} / \Delta_{32} \quad (11)$$

In code verification, S often represents point-by-point solution values or an appropriate norm of the solution error. However, in simulation verification, S typically represents a solution functional, e.g., lift coefficient. The metrics for simulation verification are sometimes called Figures of Merit (FoMs) or Quantities of Interest (QoIs) and determining their values is typically the reason for conducting the set of calculations. Solution functionals involving discontinuous features, such as shock waves, may be problematic. Baurle demonstrated simulation verification for shock boundary layer interaction within a scramjet isolator, and circumvented the issue of discontinuous properties by selecting integrated surface force components and the incipient shock location as solution functionals rather than local measures of flow properties in the vicinity of the shock waves.²⁶

If $0 < D < 1$ then monotonic convergence is obtained and RE is employed to evaluate the convergence rate (or estimated order of accuracy) p_{RE} and the error estimate δ_{RE} :

$$p_{RE} = \ln(\Delta_{32} / \Delta_{21}) / \ln(r) \quad (12)$$

$$\delta_{RE} = \Delta_{21} / (r^{p_{RE}} - 1) \quad (13)$$

The numerical uncertainty, U_{NUM} , is defined as the estimate of discretization error such that the interval $\pm U_{NUM}$ bounds the true value of δ at a 95% confidence level and is expressed as

^{††} A three-level RE is also known as Aitken's extrapolation.

$$U_{NUM} = FS|\delta_{RE}| \quad (14)$$

where FS is the verification method's Factor of Safety. In the Grid Convergence Index method, Roache suggests $FS=1.25$ with a minimum of three grids to determine p_{RE} and δ_{RE} and $FS=3.0$ for two-grid sensitivity studies using the theoretical convergence rate, p_{TH} .²⁷ When the solutions are in the asymptotic range p_{RE} is equal to p_{TH} . The ratio

$$P = \frac{p_{RE}}{p_{TH}} \quad (15)$$

is a metric for the distance of the solutions from the asymptotic range. A key element of verification is evaluating the rate of convergence to estimate how the errors would decrease as the mesh is further refined.

Xing evaluated four estimates for FS by comparing error estimates from 17 studies with analytical and/or numerical benchmark solutions.²⁸ Xing also proposed a Factor of Safety that is a function of P :

$$U_{NUM} = FS(P)|\delta_{RE}| = \begin{cases} [FS_1P + FS_0(1 - P)]|\delta_{RE}|, & 0 < P \leq 1 \\ [FS_1P + FS_2(P - 1)]|\delta_{RE}|, & P > 1 \end{cases} \quad (16)$$

Statistical analysis of a large number of samples based on the analytical and numerical benchmark solutions yielded the recommended values: $FS_0=2.45$, $FS_1=1.6$, and $FS_2=14.8$ and thus the expression for U_{NUM} given by Eq. (16) becomes:

$$U_{NUM} = FS(P)|\delta_{RE}| = \begin{cases} [(2.45 - 0.85P)]|\delta_{RE}|, & 0 < P \leq 1 \\ [(16.4P - 14.8)]|\delta_{RE}|, & P > 1 \end{cases} \quad (17)$$

When the sequence of solutions is either oscillatory or divergent ($D < 0$ or $D > 1$), Rider, et al. observe that an imprecise estimate of discretization is better than no estimate.²² In this case, they recommend the following heuristic for U_{NUM} :

$$U_{NUM} = 3(S_{max} - S_{min}) \quad (18)$$

When an uncertainty estimate for the iterative solver error, U_{TOL} , is determined then the overall uncertainty from the numerical approximation, U_{APPROX} , is obtained from the RMS of the two estimates:

$$U_{APPROX} = \sqrt{U_{NUM}^2 + U_{TOL}^2} \quad (19)$$

Richardson Extrapolation (RE) techniques may be applied to an unstructured base grid, if a systematic method of coarsening or refinement is employed from the base grid. However, when the refinement is also unstructured then r is not defined. An effective r that is based on the number of elements or vertices of the coarse and fine grids can be used but it lacks the firm basis of structured grid refinement and may significantly over- or under-estimate the discretization error.

The RE technique works well for the integral parameters. It is problematic when applied to local variables. Grid refinement needs to be conducted at least three times and preferably four times. It is difficult to identify if and when the asymptotic range of the grid convergence is reached; the technique does not work for oscillatory grid convergence.²⁹ It does not provide information about how to adapt mesh to reduce the discretization error in the quantity of interest.

B. Error Transport Equation Method

Another approach for solution verification that is more naturally suited to unstructured meshes is error estimation where error equations are derived and solved simultaneously with the original partial differential equations (PDE). At present, the ETE is based upon the upwind construction of the computational scheme and substitution of the discrete solution into the original PDE. The right hand side of the resulting modified equation represents the truncation error, the difference between the original PDE and the discretized equation. These terms form the source terms of the ETE. Truncation errors generated in one area of the

domain are convected and manifested elsewhere as interpolation errors. For hypersonic systems, errors arising from numerical discretization may act as erroneous waves that the equations propagate like physical waves. Mesh adaptation that is tied to sources of discretization errors promises to be more efficient than adaptation to local errors alone.

Cavallo extended the work of Zhang, who applied the approach to the 2D Euler equations that were solved with Roe's explicit upwind flux differencing scheme.^{23,24} Cavallo developed approximations to the truncation error for viscous fluxes by decomposing the error into two components: a) error in velocity (assuming that viscosity is correct) and b) error in viscosity (assuming that velocity is correct). An example of the ETE approach for an axisymmetric supersonic turbojet exhaust into still air is illustrated in Figs. 5-6. Three structured grids were employed with 13,301, 36,308, and 80,590 vertices, respectively. The coarse grid and the non-dimensional velocity obtained from the fine grid solution are shown in Fig. 5. The velocity field reveals the shock cells characteristic of supersonic jets.

The centerline velocity distributions for the coarse and fine grid solutions are shown in Fig. 6. Numerical error bars are also shown as defined by $u \pm |\varepsilon_u|$ where ε_u is the calculated error in velocity. Richardson extrapolation was also performed where the solution sequence exhibited monotonic behavior and is presented. The viscous ETE is found often to capture the difference between the coarse grid and extrapolated results.

The ETE approach shows promise for adaptive meshing by identifying locations where discretization errors originate. However, the method does not provide *how* to efficiently adapt in these locations. Additionally, demonstration of the ETE approach for diffusion dominated flows and the extension to other physical phenomena like chemical kinetics is needed.

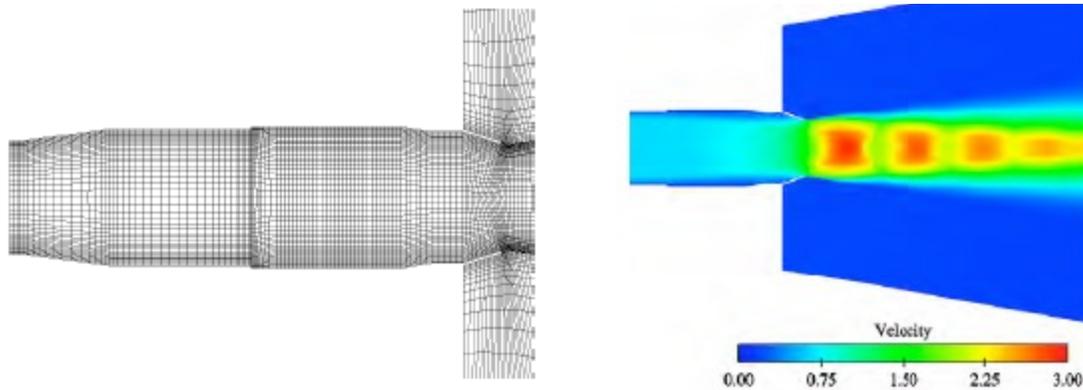


Figure 5. Near field of aircraft exhaust grid (left) and fine grid velocity field solution (right) from Ref. 23.

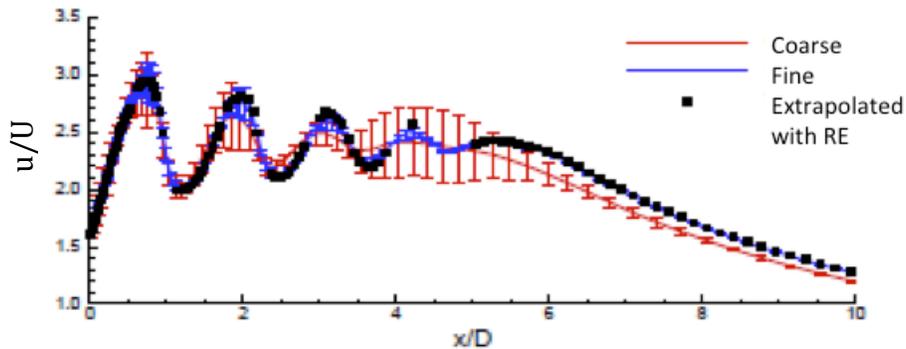


Figure 6. Predicted errors in centerline engine exhaust velocity (adapted from Ref. 23).

C. Automatic Verification of Goal-Oriented Flow Simulations

The objective of most engineering applications is the prediction of specific output quantities, e.g., the coefficients of lift and drag. The motivation for goal-oriented verification is to reliably predict the level of discretization error in these target outputs. Moreover, once the error is known, this information can be used to guide adaptive cell refinement to reduce the error below a user-specified tolerance.

Nemec and Aftosmis demonstrate the utility of the combination of an adjoint-based error estimate with a robust mesh generation tool based on a Cartesian cut-cell method to yield automated simulation verification for inviscid, compressible flow fields.²⁵ The approach is motivated by considering an output functional $J(Q)$ determined from a flow solution Q that satisfies the flow equations:

$$R(Q) = 0 \quad (20)$$

Employing a truncated Taylor series expansion about the functional and the residual equations yields

$$J_h(Q_h) \approx J_h(Q_h^H) - \psi_h^T R_h(Q_h^H) \quad (21)$$

where Q_H , Q_h , and Q_h^H represent the flow solution on the working mesh with characteristic cell-size H , the flow solution on the next uniformly refined mesh (i.e., embedded mesh) with characteristic cell-size h , and the reconstruction of the flow state from the working mesh to the embedded mesh via a prolongation operator, $Q_h^H = P Q_H$. The adjoint variable ψ is obtained via the solution of the following linear system:

$$\left[\frac{\partial R_h(Q_h^H)}{\partial Q_h} \right]^T \psi_h = \frac{\partial J_h(Q_h^H)}{\partial Q_h} \quad (22)$$

The implementation splits the adjoint-weighted residual term in Eq. (21) into a functional correction term:

$$J_C = J_h(Q_L) - \psi_{TQ}^T R_h(Q_L) \quad (23)$$

and an estimate of the remaining error in each cell of the working mesh:

$$\eta_H = \sum_{j \in V_i} (\psi_{TQ} - \psi_{TL})^T R_h(Q_L)_j \quad (24)$$

where ψ_{TL} and ψ_{TQ} are tri-linear and tri-quadratic polynomial prolongations of ψ_H and j denotes the j th child of the parent cell V_i on the working mesh. An expression for an estimate of the discretization error is:

$$\varepsilon = C |J_C - J_H(Q_H)| \quad (25)$$

where C is a constant dependent on the order of discretization, e.g., $C = 4/3$ for second order schemes. The approach controls the magnitude of η_H via mesh adaptation and uses Eq. (25) to estimate the level of discretization error in the output.

The principle of error equidistribution is used to control the cell size as the simulation advances. That is, cells are refined to make $|\eta|_H$ uniform and small, so that each cell contributes equally to improving the accuracy of the simulation in the limit. Nemec and Aftosmis found $|\eta|_H$ to reliably identify critical regions of the mesh for adaptation and Eq. (25) to provide a sharp estimate of discretization error. The CPU cost per iteration for solving the adjoint equations was reported to be approximately equal to the cost of solving the flow equations.

An important feature of this approach is that convergence histories of the outputs and error estimates are available automatically. This feature is illustrated here with the prediction of the near-field pressure signature of a lifting wing-body model in supersonic flow (Fig. 7). Figure 8 presents convergence of the signature, the functional (line integral of the signature), and its error estimates. Figure 8(a) shows that after 12 adaptations, the pressure signature is essentially mesh converged and agrees well with experimental data. Figure 8(d) combines information presented in Figs. 8(b) and 8(c) to precisely demonstrate simulation verification and asymptotic convergence.

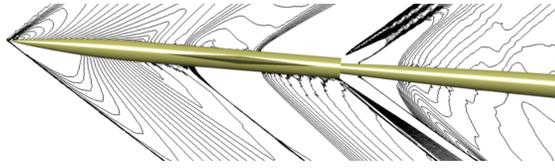
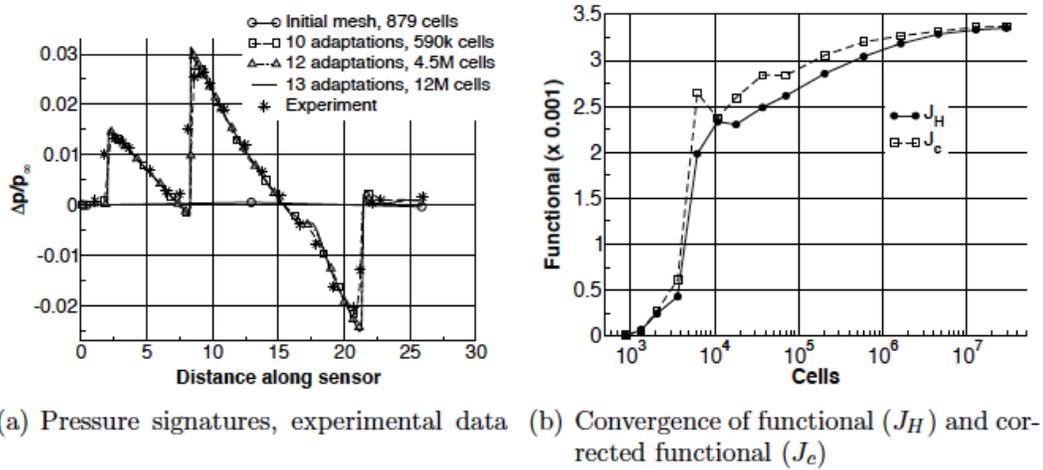
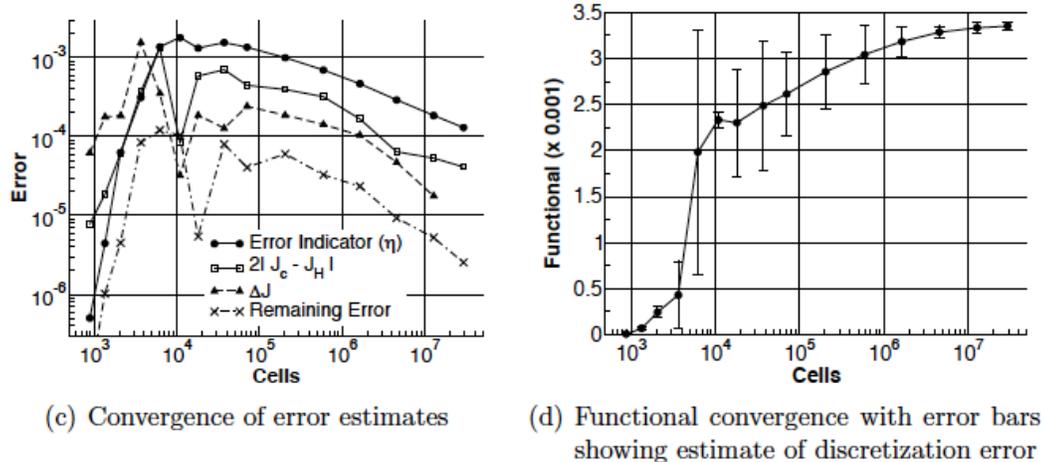


Figure 7. Symmetry plane, pressure isobars in the range 0.65-0.78 in 0.005 increments at $M_\infty = 1.68$ near a wing-body model from Ref. 25.



(a) Pressure signatures, experimental data (b) Convergence of functional (J_H) and corrected functional (J_c)



(c) Convergence of error estimates (d) Functional convergence with error bars showing estimate of discretization error

Figure 8. Mesh convergence for the pressure signature, the functional and its error estimates from Ref. 25.

Simulation verification of the prediction of aerodynamic forces for the complex problem of launch-abort-vehicle prototype presents another example of the capability for the approach to provide convergence histories of the outputs and error estimate automatically. This case considers a high-altitude abort at $M_\infty = 4$ and $\alpha = 20^\circ$ involving strong Attitude Control Motors near the nose (Fig. 9). The quantity of interest, J , is a weighted sum of normal and axial force coefficients, C_N and C_A , respectively. The maximum number of computational cells is limited to 50 million cells.

Figure 10 presents the convergence of the functional, the error estimates, and the aerodynamic forces. Figure 10(a) shows the value of the functional (J) on each mesh adaptation, with error bars representing the level of discretization error. The inset in Fig. 10(a) shows that over the last three adaptations the error bars bracket the functional and the changes in the functional are less than 1%. All error measures decrease after the mesh size is 2 million cells [Fig. 10(b)]. There are virtually no changes in the axial and normal force over the last two adaptations [Fig. 10(c)].

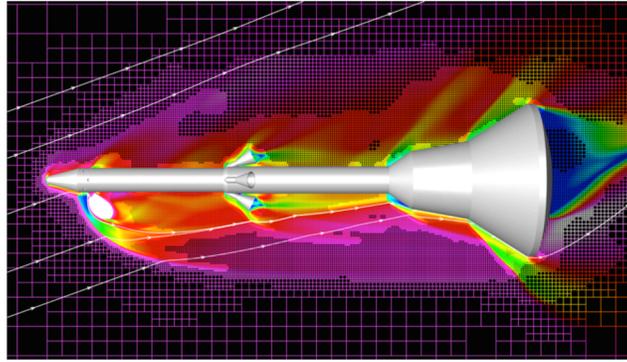
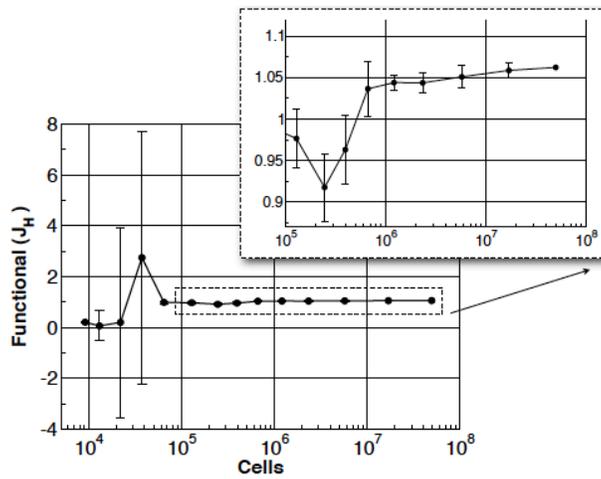
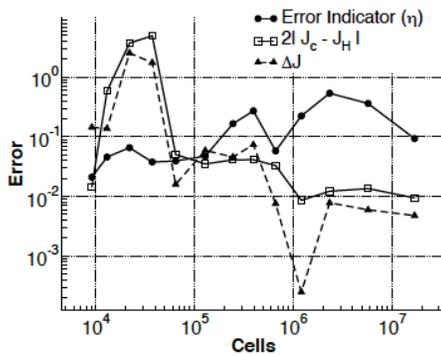


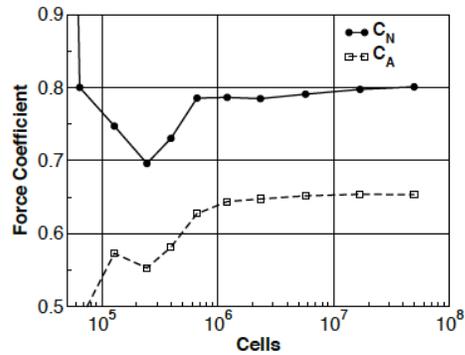
Figure 9. Mach contours after 12 adaptations on symmetry plane. Contours are compressed using \sqrt{M} to improve jet visualization (Ref. 25).



(a) Functional convergence with error bars showing estimate of discretization error



(b) Convergence of error estimates



(c) Convergence of normal and axial force coefficients

Figure 10. Mesh convergence (Ref. 25).

Overall, these examples characterize the performance and benefits of adaptive mesh refinement with adjoint-based error estimates for verification of engineering simulations for integral output quantities in an affordable and automatic manner. The fact that this approach is so powerful provides tremendous motivation for research to extend its utility to other physics simulation areas, such as unsteady high Reynolds number flows and other outputs of interest (e.g., maximum norms).

IV. Sensitivity Analysis and Propagation and Quantification of Uncertainty

“[E]xperience and instinct are poor substitutes for careful analysis of uncertainty.”^{§§}

—Return to Flight Task Group, NASA

In the Guide, Swiler and Romero provide a survey of advanced probabilistic uncertainty propagation and sensitivity analysis methods.³⁰ Specifically, the focus is on sampling design, sensitivity analysis, and response model approximations. DeLoach summarizes some basic tenants of modern Design of Experiments (DoE) that can be applied to simulation experiments and to physics experiments.³¹ Barth discusses the calculation of moment statistics with error bounds using non-intrusive input uncertainty propagation and output uncertainty quantification methods and illustrates these for computational fluid dynamic (CFD) simulations containing random parameters and fields.³²

A. Some Basic Concepts

The following definition of uncertainty is used to identify sources of uncertainty: “*A broad and general term used to describe an imperfect state of knowledge or a variability resulting from a variety of factors including, but not limited to, lack of knowledge, applicability of information, physical variation, randomness or stochastic behavior, indeterminacy, judgment, and approximation.*”³³

Uncertainty quantification (UQ) is defined as: *the process of identifying and characterizing relevant sources of uncertainties in a model, simulation, or test, and of quantifying contributions of significant sources to the total uncertainty in simulated or test quantities of interest (QoIs).* The latter quantities may be actual simulated values, measured data, or quantities derived from these values or data.

Sensitivity analysis (SA) is *the determination of how much uncertainty an individual source contributes to the total uncertainty in a simulated or test QoI.* Then the magnitudes of the individual contributions can be compared to determine the sources that primarily contribute to the total uncertainty of the test data or simulation.

Table 2 summarizes various commonly encountered *sources* of uncertainty in experiments and model simulations. Uncertainties are also characterized into various *types* according to their nature in the context of a given uncertainty analysis to identify the implication and treatment of these uncertainties in an analysis.

Table 2. Some Sources of Uncertainty in Tests, Models, and Simulations

Tests and Test Data

- **Measurement uncertainties** from sensor and data acquisition system inaccuracies, etc., in measurement of:
 - Test conditions and inputs, e.g. initial and boundary conditions.
 - System responses, behaviors, outputs.
- **Extrapolation-error uncertainty** when bias corrections and uncertainty characterizations are extrapolated from where the corrections and characterizations were quantified originally.
- **Random variability over replicate tests**, such as unit-to-unit geometric and physical variability in different tests and variability of measurement errors in different tests.
- **Data processing and inference uncertainties**, e.g.:
 - Uncertain error in temporal and/or spatial field point-data interpolation, integration, trend extrapolation, etc.
 - Uncertain error in inferring the full population of a random quantity (e.g., a frequency distribution or probability density) from limited data samples.
 - Uncertain error in uncertainty propagation and aggregation procedures that combine data measurement, processing, and inference uncertainties.

Models and Simulations

- **Uncertain values for model input parameters, or multiple values for input parameters**—e.g., populations of values described by frequency distributions.

^{§§} A corollary is the following: *Best practices are poor substitutes for careful analysis of uncertainty.*

- **Model-form discrete-choice uncertainties**, e.g., employing multiple plausible turbulence models that do not morph parametrically from one to another by varying parameter values. Other examples are multiple equation-of-state models, material constitutive models, and geometry models (model A w/bolts vs. B w/none).
- **Model prediction-bias uncertainty**: whatever the model form and parameter values, approximation error will result in some prediction bias error even if the mathematical equations are solved exactly and initial & boundary conditions inputs to the model are exact.
- **Numerical uncertainty** associated with spatial and temporal discretizations and incomplete iterative convergence of the discretized equations being solved.
- **Extrapolation uncertainty** exists when models are used away from validation points in the model validation domain or when bias corrections and uncertainty characterizations are extrapolated from where originally quantified.
- **Data processing and inference uncertainties**, e.g.:
 - Uncertainty in temporal and/or spatial field point-values interpolation, integration, trend extrapolation, etc.
 - Uncertainty in sampling, uncertainty propagation and aggregation procedures that combine the modeling and simulation uncertainties above.

Figure 11 presents a high-level diagram of some important aspects of uncertainty quantification and sensitivity analysis that can arise in analysis projects. These aspects of UQ and SA are defined and discussed briefly in Ref. 34. The aim of this diagram and associated chapter is not to be exhaustive, but to introduce aspects of UQ and SA that are often encountered in engineering practice.

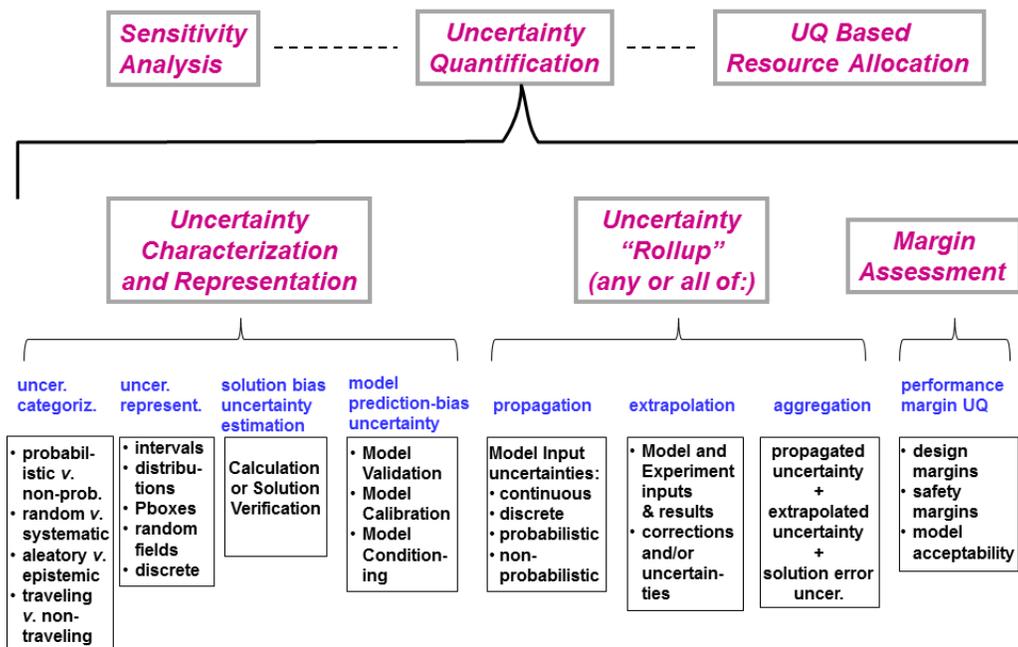


Figure 11. Some important aspects of uncertainty quantification.***

*** The terms *systematic uncertainty* and *random uncertainty* in the box labeled "Uncertainty Characterization" are not strictly proper. Paraphrasing from the GUM: the concept of *uncertainty* varying randomly from test to test is perplexing, perhaps even nonsensical. It is measurement errors themselves, or physical quantities themselves, that are conceived as randomly varying over a set of units or events, or systematically having the same value in the multiple units or events.¹¹ Nonetheless, these terms are commonplace in engineering literature and practice.

The top row of the diagram shows a linking between uncertainty quantification, sensitivity analysis, and UQ-based resource allocation. UQ-based resource allocation is briefly discussed in broad terms in Ref. 34. Some aspects of resource allocation are addressed in Ref. 32, as summarized in Section IV.B.

A. Methods and Examples

Swiler and Romero present some contemporary uncertainty propagation methods for probabilistic uncertainty in computational applications, including sampling methods, response-surface surrogate model based methods, and optimization-based reliability methods (Fig. 12).³⁰ Specifically, they discuss Monte Carlo sampling methods, Latin Hypercube sampling (LHS), and quadrature methods. Two approaches for sensitivity analysis are discussed: correlation and variance-based indices. They also demonstrate sampling, quadrature, sensitivity analysis, use of surrogates, and reliability analysis on a cantilever beam displacement problem.

These advanced propagation methods complement the classical and modern experimental design probabilistic UQ approaches presented by DeLoach, and the linearized probabilistic UQ approaches employed for validation by Romero, et al.^{31,35} Propagation of probabilistic discrete uncertainties and non-probabilistic interval uncertainties are also demonstrated in Ref. 35.

Techniques of DoE are important for efficient extraction of system behavior, performance, and sensitivity information from experiments. Validation for the intended use of simulations and models and model calibration are efficiently and properly done if DoE is used. DeLoach's objective is to provide a general roadmap to DoE.³¹

A distinction is made between "conventional" and "designed" experiments. The former is described as "data collection." Experimentation differs from data collection in that its objective is to enable the *creation* of a database to facilitate inferences about a system of interest. This objective is achieved by developing mathematical models that adequately predict responses of interest over specified ranges of selected factors upon which the responses depend.

DeLoach discusses variance in experiments, scaling, site selection in design space, response model formulation, response model adequacy assessment, and error propagation theory.³¹ A few case studies are presented to demonstrate the techniques.

Barth illustrates advanced quadrature-based methods for propagating probabilistic uncertainty in CFD.³² Uncertainty statistics for fluid dynamic quantities of interest are often estimated from a sequence of simulations that account for sources of uncertainty in a simulation model. Unfortunately, the accuracy of these estimated statistics might be severely compromised by the presence of numerical errors of two types: (1) numerical errors occurring in simulations and (2) numerical errors occurring in the calculation of uncertainty statistics integrals using a numerical method. A major challenge for simulation practitioners is determining the level of accuracy that must be achieved in simulations and the numerical calculation of statistics integrals so that output statistics are estimated as accurately as possible for a given computing cost.

A primary objective is to construct computable error bound formulas for systematically determining how accurately simulations should be computed and how accurately uncertainty statistics integrals should be approximated to attain efficient and accurate moment statistics for output quantities of interest. Formal error bounds are presented for uncertainty statistics that are estimated using simulations containing numerical errors together with numerically approximated statistics integrals with quadrature error. These error bound formulas provide a quantitative guide when performing simulations by answering the following questions:

- How accurate is a computed moment statistic?
- How does simulation numerical error affect the accuracy of a computed moment statistic?
- How does statistics integral error affect the accuracy of a computed moment statistic?
- To further improve the accuracy of computed output statistics, should additional resources be devoted to solving simulation problems more accurately or improving the accuracy of numerically approximated statistics integrals?

The ability to quantitatively answer these questions and adjust the accuracy of simulations as needed appears to be a powerful new capability in uncertainty quantification. Several numerical examples are presented in Ref. 32 to demonstrate this capability.

Approaches for Propagation of Input PDF Uncertainties

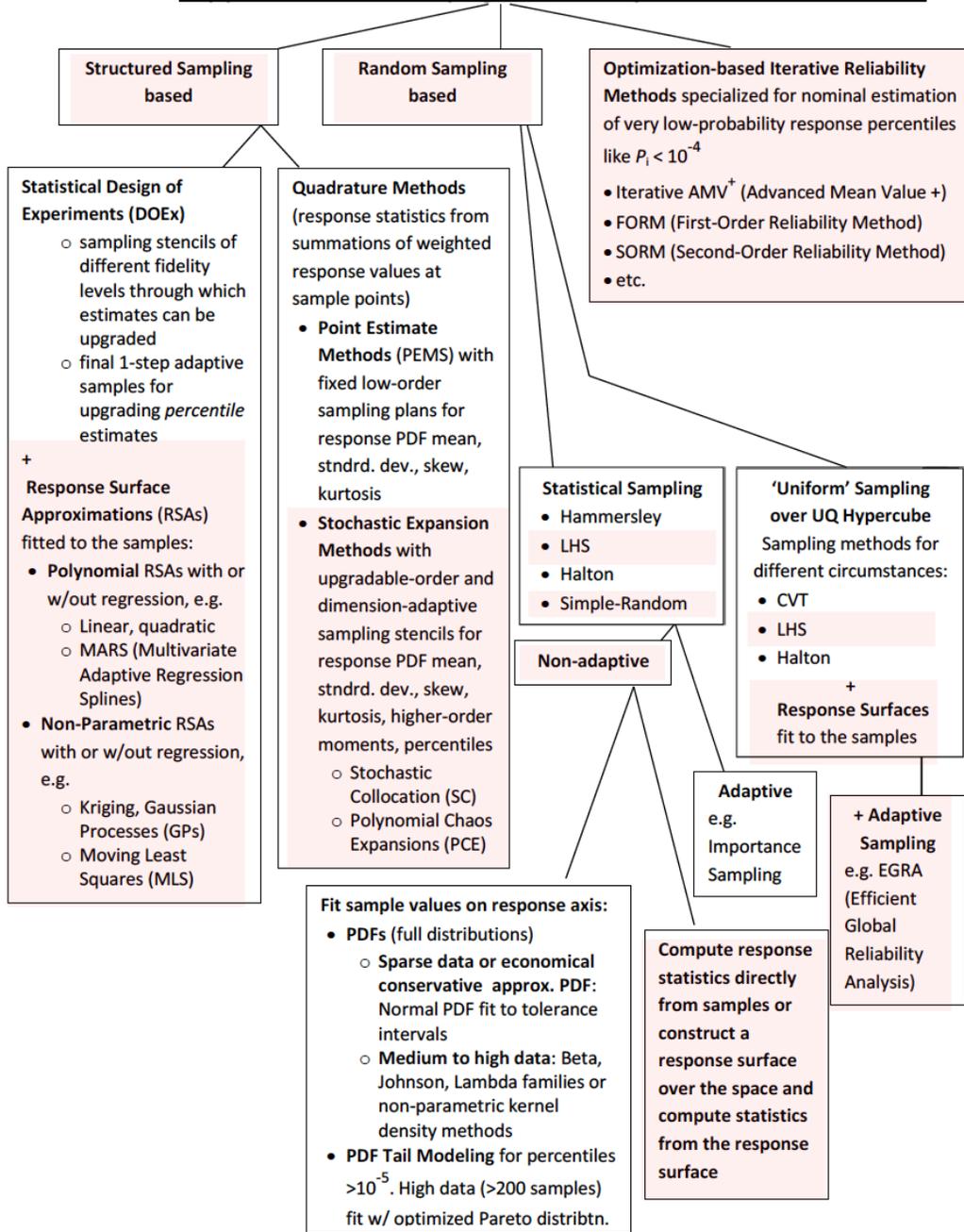


Figure 12. Propagation approaches for input PDF uncertainties highlighted in pink are described and demonstrated in Ref. 30.

One of the examples considers uncertainty propagation for a launch vehicle simulation with exhaust plume modeling (Fig. 13). Calculations are performed to quantify the extent of the Plume Induced Flow Separation (PIFS) flow reversal with respect to uncertain thrust conditions and uncertain flight Mach number (M_∞). Uncertainty in the rocket motor thrust is included in the calculation using a simplified model of thrust uncertainty given two thrust settings

$$\text{thrust}(\zeta) = \text{thrust}_{80\%} + \zeta (\text{thrust}_{100\%} - \text{thrust}_{80\%}) \quad (26)$$

where $\zeta = \text{Gaussian}_{4\sigma}$ with mean = 0.7; standard deviation $\sigma = 0.1$

$$M_\infty = \text{Gaussian}_{4\sigma} \tag{27}$$

with mean = 6.7, standard deviation $\sigma = 0.067$

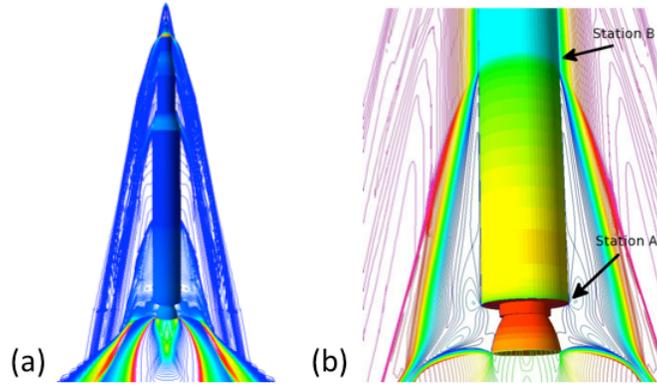


Figure 13. Mach 6.7 flow over launch vehicle configuration with rocket plume modeling.³² Shown are (a) pressure contours and (b) flow field Mach number and surface flow field density. Blue color denotes low values and red color denotes high values.

Figure 14 graphs uncertainty estimates for skin friction coefficient between Stations A and B using a 6-point (level=4) Hybrid quadrature (HYGAP) approximation with piecewise polynomial and a Clenshaw-Curtis quadrature for smooth random variable data, and cubic piecewise polynomials for non-smooth random variable data. At physical location $x = 3460$, the output probability density is bi-modal with 2 widely separate peaks. This uncertainty in separation position is the information sought by engineers.

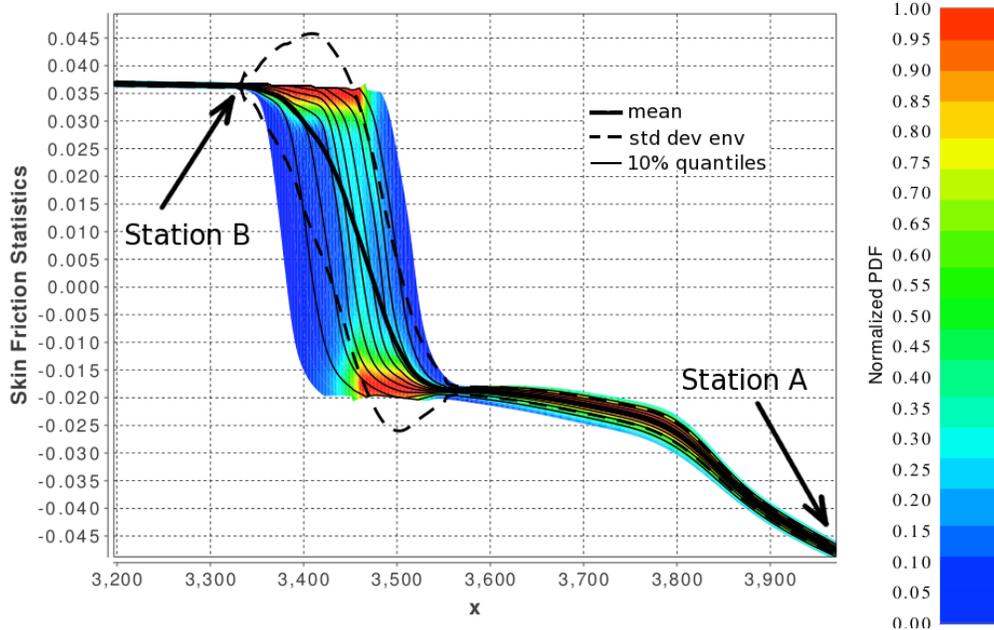


Figure 14. The skin friction coefficient statistics and 10% quantiles of probability at Mach 6.7 flow over a launch vehicle configuration.³⁶

Whether to use dense tensorization quadrature, sparse tensorization quadrature, or a multi-level (M-L) sampling method depends on uncertainty dimensions, accuracy desired, regularity (smoothness) of data, and affordability. Figure 15 presents a comparison of these methods. In practice, different approaches for the problem at hand need to be considered in order to decide which method to use. The desired accuracy

depends on the intended use of simulation. The simulation expense depends on the method used for propagation and quantification of uncertainties, and increases with the number of uncertainty dimensions and the required accuracy.

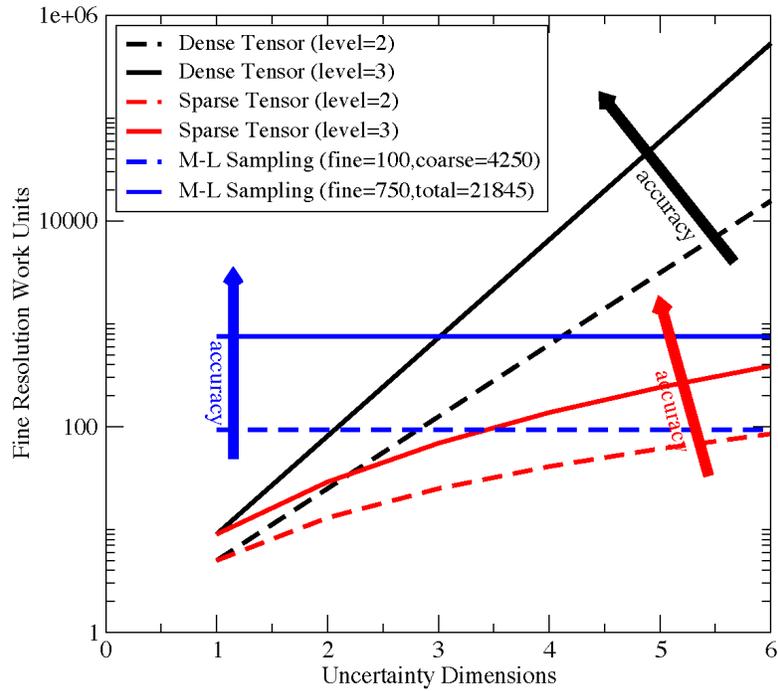


Figure 15. Quadrature complexity depicted for dense tensorization quadrature, sparse tensorization quadrature, and multi-level (M-L) sampling (courtesy of Timothy Barth).

V. Real Space Model Validation

“Nothing is as empowering as real-world validation, even if it's for failure.”

— Steven Pressfield

A. Some Initial Concepts and Paradigms

Model Validation is still a developing field in engineering and science. A large variety of viewpoints and methodologies regarding model validation exist in the literature. References such as 6, 7, 8, 9, 10, 37, 38, 39, 40, 41, 42, and 43 explore many important philosophical and implementation issues in model validation, and survey various specific paradigms and tactical approaches for performing model validation for application problems. Some of these references go on to develop a specific approach and implementation framework.

Romero, et al. use the “Real-Space” (RS) model validation approach.³⁵ The RS uncertainty accounting and comparison system for comparing experimental and model results in the context of their uncertainties is used as a model validation metric as discussed below, but is also used for model calibration.

The rationale for the Real Space validation paradigm and framework is presented in Refs. 16 and 42, and referenced papers therein. The approach adopts some elements and constructs from the literature (sometimes adding needed refinement), and adds pivotal new elements and constructs. The RS approach reflects pragmatism, versatility, and capabilities derived from working many industrial-scale problems involving complex physics and constitutive models, steady-state and time-varying nonlinear behavior and boundary conditions, and various categories of uncertainty in experiments and models. In particular, the following philosophical, conceptual, and implementation considerations underlie the formulation of the RS approach.

“Traveling Models” and “Experiment Models” and definition of ‘Model’ in Validation and Prediction Contexts

It is important to identify what model or set of models is being validated in a validation activity. Often, validation frameworks and activities are ambiguous in this respect, which can lead to confused interpretation and improper use of validation results, as well as improper accounting of uncertainties and their propagation to predictions. The RS paradigm’s concepts of *experiment models* and *traveling models* are essential concepts in distinguishing what model or set of models is being validated.^{16,42,44}

Consider a 1-D heat conduction experiment involving a heated rod. If a finite-element model is built to simulate the experiment, how does one target or differentiate whether the 1-D heat diffusion equation (partial differential equation, PDE) alone is being validated, or the validation applies to the larger set of models (equations and parameter values/ranges) consisting of the PDE and the geometry, material property, and boundary condition descriptions? The “model” to be validated could also potentially include the affiliated discretization scheme and solution algorithm.

It is necessary to be specific about which model or set of models is the focus of a validation exercise. This is important for planning and performing the validation activity and interpreting and using the results. The notion of a *traveling model* is employed to delineate the set of models in a validation or calibration activity that will be used (as a set) in subsequent predictions. This is a subset of the larger set of models employed in the validation or calibration activity, referred to as the *experiment model* or *e-model*. This larger set of models cannot be avoided. No matter what the model of traveling interest is—whether a set of PDEs, a material behavior model, a model of a hardware device, a process model, etc.—the experiment will have aspects that need to be modeled that are *auxiliary* to the traveling model of interest.

Although only a certain portion of an experiment model travels to application/prediction settings away from the validation setting, the uncertainty that issues with the traveling model is often accumulated from various uncertainties in the validation experiments, including those from the non-traveling aspects. Uncertainties in the non-traveling aspects depend on the design and implementation of the experiments, the diagnostic instrumentation used, and the chosen scope of the e-model, all of which should be optimized by good experiment design to reduce the non-traveling uncertainty as much as achievable within project constraints. Uncertainties in the experiments are treated differently in the Real Space framework if the uncertainties are affiliated with traveling aspects vs. non-traveling aspects for reasons associated with extrapolation risk as demonstrated in Refs. 16 and 44.

The traveling model can contain “*model-intrinsic*” uncertainties that are inherently affiliated with the traveling model, such as an uncertainty range on parameter values in a turbulence or material model. These uncertainties are defined prior to the current experiment (they are not determined by or in the experiment) and come to the experiment model as *a priori* uncertainties in model form and/or parameter values.

Validation Metrics and Real Space Discrepancy Characterization

Validation metrics are defined in ASME V&V 10 as mathematical measures that quantify the level of agreement between simulation and experimental outcomes.⁹ The Real Space validation approach is so named because it differs from other established frameworks that employ validation metrics in a discrepancy *transform space*, e.g., the subtractive difference metric in Refs. 10 and 12, and the CDF-mismatch ‘area’ validation metric in Refs. 39, 45, 46, and 47. Figure 16 presents some differences between these approaches. These transform measures have varying transparency and interpretability of the physical and decision-making significance of the numerical values yielded by the measures. The transforms can also constrain what forms and types of uncertainty can be handled appropriately (see Refs. 16 and 42).

Real Space discrepancy characterization better reveals characteristic differences between model and experimental results that affect prediction risk. All transform metrics appear to have non-exclusive mappings between real space and transform space, where the same transform-space value can accompany different conditions in real space. Therefore it can be risky to use transform space metrics to make validation judgments on model performance and adequacy, and to guide model conditioning and extrapolation.

Furthermore, workable approaches in the literature to assess *adequacy* of model-experiment agreement in transform space remain elusive, whereas a “zeroth order” solution to this non-trivially difficult issue can be applied in the RS approach (see case study in Section V.C, and Refs. 16 and 42 for an extensive discussion). This grew out of the necessity for transparent, intuitive, and readily interpretable results in

applied model validation projects. Importantly, the RS discrepancy characterization also segregates aleatory and epistemic uncertainties and is especially suited for assessing models to be used in the analysis of performance and safety margins (see case study example below).

Hence, if the scope of use of discrepancy characterization includes model adequacy determination, extrapolation support, and prediction risk in downstream use of the model, the Real Space approach has significant advantages over transform discrepancy characterizations like the subtraction and area transform metrics.

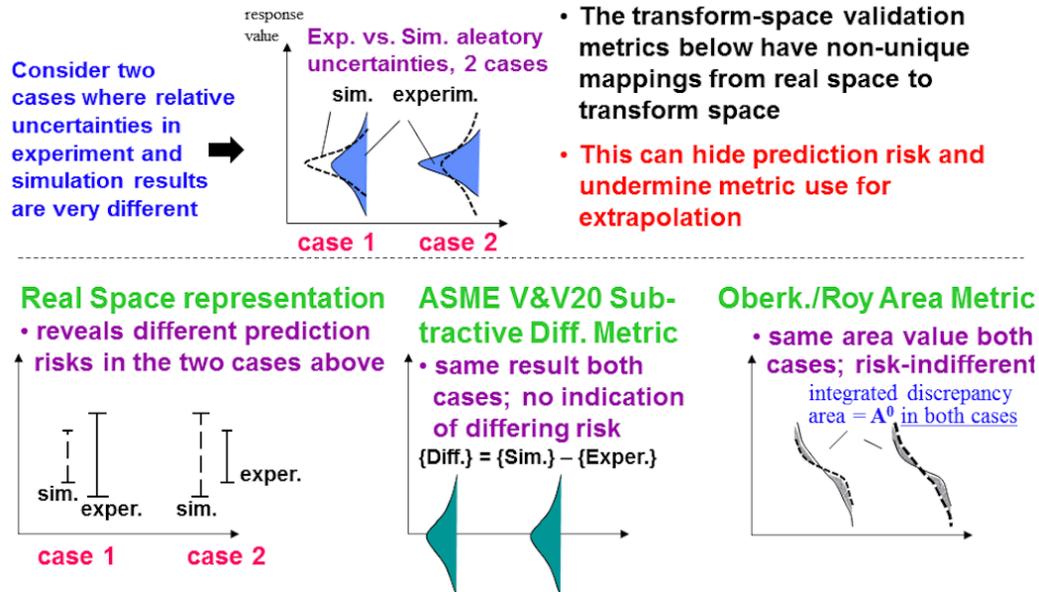


Figure 16. Real Space and transform space representations of model discrepancy (from Ref. 16 with permission).

Design of Model Validation (and Calibration) Experiments and Scope of the Experiment Model

Because of the vagaries that extrapolation presents, it is desirable to plan validation experiments as close as possible to the actual conditions of the applications for which the model is to be used. Of course this is also a tactic for calibration of models. This increases the applicability and relevance of the validation or calibration results to the intended applications.

Countervailing objectives and constraints often cause validation experiments to be relatively simple and far (in modeling parameter space) from the intended model use conditions. One driver is the need to control conditions in the validation experiments in order to maximize resolution power to isolate model bias. Cost and technical practicality also often drive validation experiments to be simpler than the eventual applications. In many cases, such as nuclear weapons testing and nuclear power plant accidents, tests near to the intended application space are not even feasible.

Moreover, the credibility of the model also depends on the judgment of its robustness in extrapolation to the end-use conditions. Extrapolation robustness, in many cases, depends on if the “integral” model provides acceptable results in the validation conditions, and whether it provides these principally “for the right reasons.” This can be established to some degree via a *hierarchical* validation approach. This decomposes the model into sub-groupings of phenomenological effects at an affordable level, which allows for the validation of the separated pieces of the model. Optimally cost-effective decomposition depends on many things. Modeling and simulation are very useful for helping to optimize the decomposition. Hierarchical model validation decomposition procedures, and information integration from the separate validations, are active areas of research.

Balanced judgment over all the considerations involved must be applied to derive the most benefit from validation experiments and assessments. Other aspects of good experiment planning include: a) the use of modeling to help design the specific experiments and diagnostic instrumentation, and b) optimizing

the scope/boundaries of the experiment model (*e-model*) and of the traveling model to reduce uncertainty and increase resolution in the validation assessment.¹⁶

B. Some Specifics of the Real Space Model Validation Approach

By working backwards from an end objective of “best estimate with uncertainty” (BEWU) modeling and prediction, the RS model validation approach was arrived at. This objective was pursued within the full realistic context of integrated experimental-modeling-analysis programs with end-to-end scope involving experiment design and analysis, data conditioning, model conditioning/calibration, model validation, hierarchical modeling, and extrapolative prediction under uncertainty. Application areas have included heat transfer, solid and structural mechanics, irradiation damage in electronics, and combustion in fluids and solids. In these environments an appreciation was gained for the constraints and difficulties in devising a feasible end-to-end methodology. The RS approach evolved accordingly to address many types and sources of uncertainties in experiments, models, and simulations. This includes: traveling and non-traveling; random and systematic and closely related categories aleatory and epistemic; correlated and uncorrelated; and interval, distributional, and combined natures. These arise in and from:

- Stochastic physical systems and associated models
- Experimental variability in repeated experiments
- Associated epistemic uncertainty from limited numbers of repeated experiments
- Measurement uncertainties in experimental inputs and outputs
- Uncertainties that arise in data processing and inference from raw experimental and simulation outputs
- Parametric and discrete-form uncertainties affiliated with the model
- Numerical solution uncertainty from model discretization effects

The following steps are pursued within the RS validation framework. The steps are thought to generically apply for other model validation frameworks as well, except for the particulars of the RS discrepancy characterization approach and associated uncertainty treatment. Other validation frameworks may contain other steps.

1. Identify the modeling setting(s), e.g., initial and boundary conditions and usage circumstances, and corresponding Quantities of Interest (QoIs) for which model predictivity is to be assessed in the validation setting(s).
2. Decide on specific validation settings and activities (identify achievable experiments and outcomes) according to the following constraints and objectives. Constraints: project resources of funding, allotted time, personnel, capabilities and expertise, experimental apparatus, “allowed” and achievable experiments, and experimental conditions. *Objectives*: Validation experiments should be as close as reasonable to the actual conditions of the applications for which the end-use model will be used. This reduces uncertainty associated with extrapolation. But the objective of establishing a degree of credibility that the model is principally “right for the right reasons” calls for a hierarchical decomposition and approach to the validation problem, which competes with resource priorities of the (also important) integral validation activities. Modeling and simulation should be used, as possible, to help optimize hierarchical decomposition (if any) of the validation problem and to design specific experiments and locations of output response measurements.
3. Perform the planned experiments and form uncertainty representations for the correlated and uncorrelated errors in measurement, inference, and estimation of the experimental inputs and outputs in the multiple replicate experiments. Then categorize respectively as either traveling or non-traveling. In each category, the uncertainty could either epistemic/systematic or aleatory/random. (If only one experiment, all uncertainty is considered correlated/epistemic/systematic in the Real Space method.)
4. Process experimental results and uncertainties according to the UQ methodology explained and demonstrated in Ref. 35. Obtain uncertainty of QoI response statistics of interest for validation comparisons to predicted results: mean behavior, variance, and/or selected percentiles of response.

5. At the reference nominal values of the experimental input conditions that anchor the processed uncertainty of experimental and simulation QoI results, perform solution verification. Or state a plausible modeling scheme to be used to attain consistent (traveling) discretization-related solution bias error in model validation and end-use settings (plausible schemes can exist in some cases). If a solution verification approach is used, perform solution verification at one or more other strategically selected points in the uncertainty parameter space (see next step) of the validation exercise. This is used to assess whether the discretization correction used in the following step is robust over the UQ parameter space. See Ref. 48 for selection of strategic points for such solution verification studies.
6. Categorize the uncertainties associated with the model into the structure in Step 3 and process the uncertainties accordingly using the UQ methodology.³⁵ Propagate uncertainties with the highest-resolution mesh and solver combination that can be afforded for the chosen UQ propagation procedure. Often, construction of a linear or low-order polynomial response-surface surrogate model UQ approach is sufficient for model validation UQ purposes. From the propagated and aggregated uncertainties, bias-corrected from the information from step 5 for the particular mesh/solver used, obtain uncertainty of QoI response mean, variance, and/or selected percentiles for validation comparison to corresponding experimental results.
7. Compare uncertainties of experimental and predicted statistics of QoIs. Quantify prediction bias uncertainties for these statistics for each QoI. See end of subsection C for the manner of comparison and interpretation of results in the Real Space approach.

C. Illustration of Model Calibration, Verification, and Validation in Case Study Example

A case study involving model calibration, verification, validation, and associated uncertainty quantification involving heated pipes pressurized to failure is presented in Ref. 35. The case study involves:

- Material strength characterization and associated constitutive model calibration.
- Model-assisted design and analysis of the validation experiment to characterize and reduce uncertainty.
- Mesh and solver discretization studies (simulation verification) to control and characterize solution error and uncertainty.
- Model validation comparison of experimental and simulation results with uncertainties and interpretation of the results.

The project's finite-element (FE) models, geometries, mesh and solver choices, and calculation verification studies to characterize discretization related solution errors are described in Ref. 35. The FE model and simulations employed in the constitutive model material characterization/inversion procedure emulate the cylinder test specimens' response, through deformation ("necking") and failure in the tension tests. Various versions of the heated pressurized pipe models and simulations employed are described. The use of modeling and simulation is demonstrated to help design the pipe-level validation experiments and thermocouple locations to minimize errors and uncertainty associated with the experiments and modeling of the boundary conditions from spatially sparse sensor information.

The validation experiments and simulations, their uncertainties, and processing of results and uncertainties for comparison in the RS framework are described.³⁵ Relatively simple spreadsheet calculation procedures are explained for the linearized UQ used. Methods for representing and propagating interval and probabilistic uncertainties from multiple correlated and uncorrelated sources in the experiments and simulations are demonstrated for: a) material variability characterized by non-parametric random functions (discrete temperature dependent stress-strain curves); b) very limited (sparse) experimental data at the coupon testing level for material characterization and at the pipe-test validation level; c) boundary condition reconstruction uncertainties from spatially sparse sensor data; d) normalization of pipe experimental responses for measured input-condition differences among replicate tests and for random and systematic uncertainties in measurement/processing/inference of experimental inputs and outputs; and e) numerical solution uncertainty from model discretization and solver effects.

Among the many model validation paradigms, frameworks, and methodologies in the literature, the Real Space methodology was found to be uniquely capable to handle the diverse set of challenging attributes in this problem.

Ultimately, uncertainty ranges of experimental and predicted 0.025 and 0.975 percentiles of response (failure pressure) are compared. Figure 17 presents the results. The red and green intervals at left in Fig. 17 reflect the validation-processed epistemic uncertainties of model-predicted 0.025 and 0.975 percentiles of failure pressure for an essentially infinite population of heated pressurized pipes. The red and green intervals to the right of the dividing line in Fig. 17 reflect the corresponding epistemic uncertainties of experimentally inferred 0.025 and 0.975 percentiles of failure pressure.

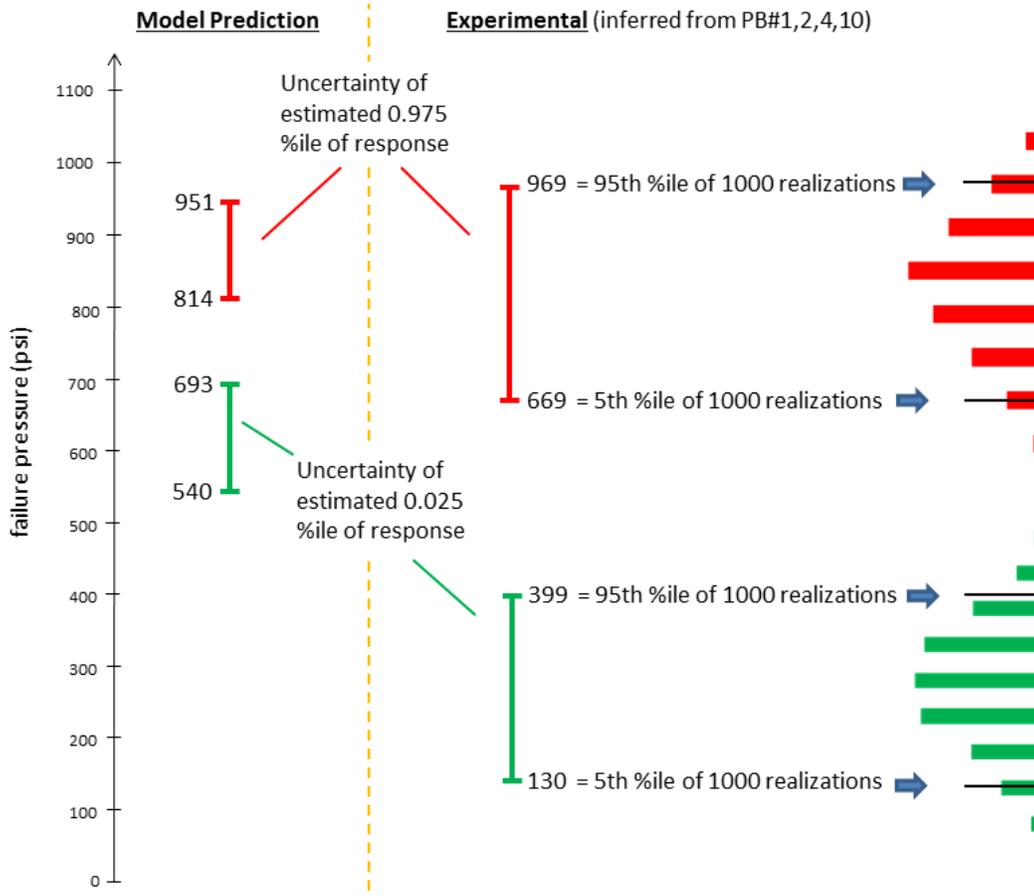


Figure 17. Model validation comparison of uncertainty ranges of simulation percentiles of response and normalized experimental percentiles of response.³⁵

Consider the green intervals for 0.025 percentiles of predicted and experimental response. The green intervals do not overlap. Therefore it is straightforward that, for this percentile of response, the model predicts higher failure pressures than inferred from testing. If, for instance, this lower percentile of response is written into a design or safety spec such that < 2.5% of pipes of this design are to fail under applied pressure and temperature conditions emulated in the tests, then the experiments are indicating a lower failure pressure for 2.5% of pipes than the model is predicting. The model therefore gives non-conservative predictions for these circumstances.

This quantitative characterization of the model's non-conservativeness for predicting the 0.025 percentile of the failure population can be used to potentially adjust or bias-correct the model for predicting this quantity, with uncertainty on the correction. Methods for this are an active area of research.^{44,49}

Indeed, robust extrapolation of the validation results in Fig. 17 to other applications of the constitutive model (different pressure vessel geometries, heating conditions, wall thicknesses, etc.) is a very difficult issue. But with some reasonable assumptions and a little more analysis one could take the results in Fig. 17 and extend them to cases where the same pipe design and experimental conditions exist but the spec has lower allowable percentages of failure like 1% or 0.001%. We would tentatively conclude that the model would be non-conservative for those spec regimes as well. With less assumptions one could reprocess the

experimental and simulation data for a more definitive assessment of prediction conservatism or not, and by how much, for specified individual percentiles of behavior. For example, the data could be reprocessed for two-sided confidence intervals on an individual percentile of interest like the 99th percentile of behavior, or could be reprocessed for one-sided confidence or tolerance *bounds* associated with a prescribed statistical confidence that a particular percentile of response meets, exceeds, or does not encroach upon a threshold response level. Such processing gets into the realm of QMU (quantification of margins and uncertainty) techniques and is beyond the scope of the present validation activity.

Now consider the red intervals in Fig. 17 for 0.975 percentiles of response. These intervals overlap and the experimental and simulation uncertainties they represent are statistically independent. Therefore there are numerous possibilities that the experimental 0.975 percentile of response is higher than the predicted 0.975 percentile, and vice versa. So we cannot conclude, as was done for the 0.025 percentile, that the predictions are non-conservative (or alternatively that they are conservative). One limiting case for these ranges of uncertainty is that the predicted 0.975 percentile is as high as 951 psi, as labeled in the figure, while the experimental percentile is as low as 669 psi, as labeled. In this limiting case the predicted 0.975 percentile is up to 282 psi higher than the experimental percentile. The opposite limiting possibility is that the predicted 0.975 percentile is as low as 814 psi, while the experimental percentile is as high as 969 psi. Then the predicted 0.975 percentile is as much as 155 psi lower than the experimental percentile. Per the discussion above, this information can be used to guide potential model bias correction for predicting this percentile of response. Processing can also be done to compare means and/or variances of the experimental and simulation results, including the uncertainties on these statistics. But processing to compare percentiles as shown here reflects uncertainty in both mean and variance and is more relevant for validation assessment of models to be used in the analysis of performance and safety margins.

The kinds of information presented here are available from model validation to help contextualize model predictiveness for decision makers. Quantitative results regarding model adequacy often come in the form of whether or not the model yields conservative predictions in the validation assessment. This is referred to as “zeroth-order” adequacy assessment in Ref. 16. Very frequently in model validation projects, any further quantitative statements on the adequacy of the model for anticipated uses are not rigorously supported or allowed by the information in the validation activity.

Nonetheless, decision makers can use the quantitative validation information of the kinds presented here to supplement other project information and considerations (quantitative and qualitative) to make more informed (although perhaps still highly subjective) decisions on model acceptability for identified use purposes. Thus, the result of a validation assessment is rarely a binary statement about whether a model is valid for a proposed use. Instead, it provides a crucial source of information in a larger overall PCMM-type assessment of suitability for the proposed use.²⁰ For use purposes that involve significant extrapolation, subject-matter experts are vital to make any judgments of model adequacy for the proposed predictions. If predictions are made with the model, the prediction bias and uncertainty information from validation can be used to nominally bias-correct extrapolative predictions, e.g., as demonstrated in Refs. 44 and 49.

VI. Takeaway

***“What we observe is not nature itself,
but nature exposed to our method of questioning.”***

—Werner Karl Heisenberg

The JANNAF Simulation Credibility Guide presents some state of the art philosophies, principles, frameworks, approaches, and processes for establishing simulation credibility. Some background information on verification and validation is briefly discussed to highlight commonality and differences and to focus on quantification of simulation uncertainty. A workflow for guiding simulation verification is presented. Three methods for conducting simulation verification are discussed: Richardson Extrapolation, Error Transport Equation, and Adjoint-Based Error Estimation. The last, in particular, is a very powerful approach when applicable. Some input uncertainty propagation methods leading to the quantification of uncertainty in output are discussed. An advanced quadrature-based method with computable error bound estimates for propagating probabilistic uncertainty is illustrated. The Real Space model validation approach is briefly discussed. Some basic tenants of modern Design of Experiments are presented in the Guide.

The presented different methods for simulation verification and for propagation and quantification of uncertainty have strengths and weaknesses. These must be considered for deciding which method to use. Output probability distributions, moment statistics, or both could be of interest. Simulation verification is relatively easier than simulation validation. Uncertainty propagation leading to uncertainty quantification is relatively easier than prediction, input parameter, or model uncertainty quantification.

Simulation validation includes simulation model validation and prediction validation. The validation of a real-world model must be done with all necessary quality experimental/test data. Nevertheless, frameworks, techniques, processes, etc. used for validation are valuable and appropriate for validating any simulation with credible referent data.

Approaches, such as the ASME V&V 20 approach, for simulation model validation and that of Real Space approach differ.^{10,16} The definitions of validation are significantly different: consequently, validation metrics are different. The ASME approach estimates the model error. The Real Space approach focuses on model adequacy for prediction. At the most, it leads to “zeroth-order” adequacy assessment. “Adequacy” is qualitative. “Accuracy” is quantitative. Is the Real Space definition of model validation appropriate? The next edition of the JANNAF Guide needs a chapter on validation that discusses and demonstrates other approaches.

There are some operational differences among the Real Space validation approach and those presented in Refs. 10, 39 and 47. Key differences are what is the experiment-simulation comparison framework, what types of uncertainties are considered, and how they are handled. The Real Space approach seems to work for various practical problems. It estimates the model error for the aleatory component of uncertainty as well, allowing a more comprehensive and granular quantification such as error on percentiles of response.

The focus of the JANNAF Guide is on simulation credibility. The decision maker or the user of simulation is interested in predictions and this person needs to know the simulation prediction error in order to make critical decisions. The credibility of prediction is what really matters. How to assess this with circumstantial evidence has yet to be properly and fully formulated for complex physics problems. “Prediction Validation” should be a new sub-category within the framework of simulation validation.

References

¹Kawaguti, M., “Numerical Solution of the NS Equations for the Flow Around a Circular Cylinder at Reynolds Number 40,” *Journal of Phy. Soc. Japan*, Vol. 8, pp. 747-757, 1953.

²Richardson, L. F., “The Approximate Arithmetical Solution by Finite Differences of Physical Problems Involving Differential Equations, with an Application to the Stresses in a Masonry Dam,” *Philosophical Transactions of the Royal Society of London*, Vol. 210A, 1910, pp. 307-357.

³Riggins, D., “Review of Physics and Modeling Issues in High-Speed Air-Breathing Propulsion,” to appear in the JANNAF Simulation Credibility Guide.

⁴Sankaran, V., Teshome, S., and Yoon, C., “Computational Modeling of Liquid Propulsion Systems,” to appear in the JANNAF Simulation Credibility Guide.

⁵Eaton, A. M., “Review Of Analytical Approaches For Modeling Solid Propulsion Rocket Motor Environments And Performance,” to appear in the JANNAF Simulation Credibility Guide.

⁶Mehta, U. B., “Simulation Credibility Level,” Presented at the 5th Joint Army-Navy-NASA-Air Force (JANNAF) Modeling and Simulation Subcommittee Meeting, Denver, Colorado, CDJSC 49, CPIAC, Johns Hopkins University, May 2007.

⁷Mehta, U. B., “Guide to Credible Computer Simulations of Fluid Flows,” *Journal of Propulsion and Power*, Vol. 12, No. 5, 1996, pp. 940-948.

⁸American Institute of Aeronautics and Astronautics (AIAA), *Guide for the Verification and Validation of Computational Fluid Dynamics Simulations*, AIAA-G-077-1998.

⁹American Society of Mechanical Engineers (ASME), *Guide for Verification and Validation in Computational Solid Mechanics*, ASME V&V 10-2006.

¹⁰American Society of Mechanical Engineers (ASME), *Standard for Verification and Validation in Computational Fluid Dynamics and Heat Transfer*, ASME V&V 20-2009.

¹¹*Guide to the Expression of Uncertainty in Measurement* (corrected and reprinted, 1995) (GUM), International Organization for Standardization, Geneva, Switzerland, 1995.

- ¹²Coleman, H. W., and Stern, F., “Uncertainties in CFD Code Validation,” *ASME J. Fluids Engineering*, Vol. 119, Dec. 1997, pp. 795–803.
- ¹³Warming, R. F., and Hyett, B. J., “The Modified Equation Approach to the Stability and Accuracy Analysis of Finite Difference Methods,” *J. Comp. Physics*, Vol. 14, (1974), 159–179.
- ¹⁴Shi, J., Zhang, Y.-T., and Shu, C.-W., “Resolution of High Order WENO Schemes for Complicated Flow Structure,” *J. Comp. Phys.*, Vol. 186, 2003.
- ¹⁵Moffat, R. J., “Describing the Uncertainties in Experimental Results,” *Experimental Thermal and Fluid Sciences*, Vol. 1, Jan. 1988, pp. 3-17. (See also Moffat, R. J., “The Measurement Chain and Validation of Experimental Measurements,” *ACTA IMEKO*, Vol. 3, No. 1, May 2014, pp. 16-18.)
- ¹⁶Romero, V. J., “Elements of a Pragmatic Approach for dealing with Bias and Uncertainty in Experiments through Predictions,” Sandia Report, SAND2011-7342, November 2011.
- ¹⁷DoD, DoD Instruction Number 5000.61, Dec. 9, 2009.
- ¹⁸National Aeronautics and Space Administration (NASA), *Standard for Models and Simulations*, NASA-STD-7009, 2008.
- ¹⁹National Aeronautics and Space Administration (NASA), *NASA Policy Directive on Technical Standards*, NPD 8070.6B, May 2003.
- ²⁰Hills, R. G., Witkowski, Rider, W. J., Trucano, T. G., and Urbina, A., “Development of a Fourth Generation Predictive Capability Maturity Model,” Sandia Report SAND2013-8051.
- ²¹Floyd, R. W., “Nondeterministic Algorithms,” *Journal of the ACM*, Vol. 14, No. 4, Oct. 1967, pp. 636–644.
- ²²Rider, W. J., Kamm, J. R. and Weirs, V. G., “Procedures for Calculation Verification,” to appear in the JANNAF Simulation Credibility Guide.
- ²³Cavallo, P., “Error Transport Methods in Solution Verification,” to appear in the JANNAF Simulation Credibility Guide.
- ²⁴Zhang, X. D., Trepanier, J.-Y. and Camerero, R., “A posteriori error estimation for finite-volume solutions of hyperbolic conservation laws,” *Computer methods in applied mechanics and engineering*, Vol. 185, 2000, pp.1-19.
- ²⁵Nemec, M. and Aftosmis, M. J., “Toward Automatic Verification of Goal-Oriented Flow Simulations,” to appear in the JANNAF Simulation Credibility Guide.
- ²⁶Baurle, R. A., Middleton, T. F. and Wilson, L. G., “Reynolds-Averaged Turbulence Model Assessment for a Highly Back-Pressured Isolator Flowfield,” 2012 JANNAF CS/APS/EPSS/PSHS Joint Subcommittee Meeting, Monterey, CA.
- ²⁷Roache, P. J., *Verification and Validation in Computational Science and Engineering*, Hermosa Publishing, 1998.
- ²⁸Xing, T. and Stern, F., “Factors of Safety for Richardson Extrapolation,” *Journal of Fluids Engineering*, Vol. 132, June 2010.
- ²⁹Celik, I. B., “Critical Issues with Quantification of Discretization Uncertainty in CFD,” NETL, 2011 Workshop on Multiphase Flow Science, Airport Marriott Station Square, Pittsburgh, PA, Aug. 2011.
- ³⁰Swiler, L. P., and Romero, V. J., “A Survey of Advance Probabilistic Uncertainty Propagation and Sensitivity Analysis Methods,” to appear in the JANNAF Simulation Credibility Guide.
- ³¹DeLoach, R., “The Modern Design of Experiments,” to appear in the JANNAF Simulation Credibility Guide.
- ³²Barth, T., “On the Calculation of Moment Statistics with Error Bounds for CFD Calculations Containing Random Parameters and Fields,” to appear in the JANNAF Simulation Credibility Guide.
- ³³National Aeronautics and Space Administration (NASA), *NASA General Safety Program Requirements* (w/Change 9 dated 2/08/13), NPR 8715.3C.
- ³⁴Romero, V. J., “Uncertainty Quantification and Sensitivity Analysis—Some Fundamental Concepts, Terminology, Definitions, and Relationships” to appear in the JANNAF Simulation Credibility Guide.
- ³⁵Romero, V. J., Dempsey, J. F., and Antoun, B., “Case Study Example: Application of UQ and V&V to Experiments and Simulations of Heated Pipes Pressurized to Failure,” to appear in JANNAF Simulation Credibility Guide.
- ³⁶Barth, T., “Non-intrusive Uncertainty Propagation with Error Bounds for Conservative Laws Containing Discontinuities,” *Uncertainty Quantification in Computational Fluid Dynamics*, Eds. Bijl, H., Lucor, D., Mishra, S., and Schwab, C., Springer International Publishing, Switzerland, 2013, pp. 1-57.
- ³⁷Anderson, M. G., and Bates, P. D., editors, *Model Validation—Perspectives in Hydrological Science*, Wiley, 2001.

³⁸Roache, P. J., *Fundamentals of Verification and Validation*, Hermosa Publishers, Socorro, NM, 2009.

³⁹Oberkampf, W. L., and Roy, C. J., *Verification and Validation in Scientific Computing*, Cambridge University Press, 2010.

⁴⁰Bayarri, M., Berger, J., Rui, P., Sacks, J., Cafeo, J., Cavendish, J., Lin, C-H., Tu, J. "A Framework for Validation of Computer Models," *Technometrics*, May 2007, Vol. 49, no. 2, pp. 138 – 154.

⁴¹Babuska, I., F. Nobile, R. Tempone, "Reliability of Computational Science," *Num. Methods in Partial Differential Equations*, (2007) Vol. 23, pp. 753-784.

⁴²Romero, V. J., "Comparison of Several Model Validation Conceptions against a "Real Space" End-to-End Approach," *Soc. Automotive Engrs. International Journal of Materials and Manufacturing*, June 2011.

⁴³National Research Council, *Assessing the reliability of complex models: The mathematical and statistical foundations of verification, validation, and uncertainty quantification*, 2012.

⁴⁴Romero, V. J., "Type X and Y Errors and Data & Model Conditioning for Systematic Uncertainty in Model Calibration, Validation, and Extrapolation," SAE paper 2008-01-1368 for Society of Automotive Engineers 2008 World Congress, April 14-17, 2008, Detroit, MI.

⁴⁵Ferson, S., W. L. Oberkampf, L. Ginzburg, "Model Validation and Predictive Capability for the Thermal Challenge Problem," *Comput. Methods in Applied Mechanics and Engrng.*, Vol. 197, 2009, pp. 2408 – 2430.

⁴⁶American Society of Mechanical Engineers, *An Illustration of the Concepts of Verification and Validation in Computational Solid Mechanics*, V&V 10.1-2012, April 16, 2012.

⁴⁷Roy, C. J., and Oberkampf, W. L., "A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing," *Computer Methods in Applied Mechanics and Engineering*, 200, 2011.

⁴⁸Romero, V. J., "Model-Discretization Sizing and Calculation Verification for Multipoint Simulations over Large Parameter Spaces," paper AIAA2007-1953, 9th AIAA Non-Deterministic Methods Conference, April 23 - 26, 2007, Honolulu, HI.

⁴⁹Romero, V. J., "Validated Model? Not So Fast. The Need for Model 'Conditioning' as an Essential Addendum to Model Validation," paper AIAA-2007-1953, 9th AIAA Non-Deterministic Methods Conference, April 23 - 26, 2007, Honolulu, HI.