

AIAA-2008-2154

An Uncertainty Structure Matrix for Models and Simulations

Lawrence L. Green,^{*} Steve R. Blatnig,[†]
Michael J. Hemsch,[‡] James M. Luckring,[§] and Ram K. Tripathi,^{**}
NASA Langley Research Center, Hampton, Virginia, 23681

Software that is used for aerospace flight control and to display information to pilots and crew is expected to be correct and credible at all times. This type of software is typically developed under strict management processes, which are intended to reduce defects in the software product. However, modeling and simulation (M&S) software may exhibit varying degrees of correctness and credibility, depending on a large and complex set of factors. These factors include its intended use, the known physics and numerical approximations within the M&S, and the referent data set against which the M&S correctness is compared. The correctness and credibility of an M&S effort is closely correlated to the uncertainty management (UM) practices that are applied to the M&S effort. This paper describes an uncertainty structure matrix for M&S, which provides a set of objective descriptions for the possible states of UM practices within a given M&S effort. The columns in the uncertainty structure matrix contain UM elements or practices that are common across most M&S efforts, and the rows describe the potential levels of achievement in each of the elements. A practitioner can quickly look at the matrix to determine where an M&S effort falls based on a common set of UM practices that are described in absolute terms that can be applied to virtually any M&S effort. The matrix can also be used to plan those steps and resources that would be needed to improve the UM practices for a given M&S effort.

Nomenclature

| | | |
|------|---|--|
| CAIB | = | Columbia Accident Investigation Board |
| CDF | = | cumulative distribution function |
| DMSO | = | Defense Modeling and Simulation Office (now the Modeling and Simulation Coordination Office) |
| IV&V | = | independent verification and validation |
| LaRC | = | Langley Research Center |
| PDF | = | probability density function |
| RPG | = | recommended practices guide |
| M&S | = | modeling and simulation |
| OCE | = | Office of Chief Engineer |
| SE | = | software engineering |
| SME | = | subject matter expert |
| UM | = | uncertainty management |
| UQ | = | uncertainty quantification |

^{*} Mathematician, Space Mission Analysis Branch, Systems Analysis and Concepts Directorate, 1 North Dryden Street, Mail Stop 462, Senior Member of AIAA.

[†] Physicist, Durability, Damage Tolerance and Reliability Branch, Mail Stop 188E.

[‡] Senior Research Engineer, Configuration Aerodynamics Branch, Mail Stop 499, Associate Fellow AIAA.

[§] Senior Research Engineer, Configuration Aerodynamics Branch, Mail Stop 499, Associate Fellow AIAA.

^{**} Senior Research Scientist, Durability, Damage Tolerance and Reliability Branch, Mail Stop 188E, Associate Fellow AIAA.

I. Introduction

Aerospace software can be loosely categorized as either (1) command, control, display, and support software or (2) modeling and simulation (M&S) software. Software that is categorized in the first group is used to provide a service function to its operators. For example, the software might provide flight command and control automation and management functions for aerospace vehicles within certain sets of predefined operational conditions, or it might provide pilots, crew, and support personnel with information about the past, current, or future states of an aerospace vehicle during operations. Command and control automation functions might be required or desired, for example, when a sequence of steps must be performed repeatedly for safety reasons, when the drudgery of a tedious set of commands can be eliminated, or when the pilot cannot respond quickly enough to control the vehicle. Display software provides information by transforming the responses from various vehicle sensors into digital or analog displays for human interpretation. Support software typically enables other control or display software to access the proper information at the proper times and in the proper formats.

Command, control, display, and support software is universally expected to work correctly all the time, with no questionable or “grey” areas. Incorrect data input to the software is expected to be the only means to arrive at incorrect answers. In this software community, verification means that the software contains no errors; validation means that the software does what it was intended to do, that is, it commands or controls the vehicle, system, or subsystem as intended. By defining the service function that the software is to perform, the correctness of the software can usually be demonstrated easily and without question to a wide variety of people, including non-specialists. This type of software is typically developed under strict management processes that are intended to reduce defects in the software product.¹⁻⁹ A large number of documents have been prepared that put forth requirements, policy directives, procedures, guidelines, and best practices for software development (e.g., see, for example, attachments 1 and 2 of ref. 6). These requirements, policy directives, procedures, guidelines, and best practices for software development are generally targeted to the software engineering (SE) community, rather than the M&S community, and tend to use language more familiar to SE community than to the M&S community. Moreover, these requirements, policy directives, procedures, guidelines, and best practices for software development generally ignore a critical aspect of M&S: uncertainty. This uncertainty in M&S may arise from imperfectly modeled physics (either through intentional approximation or through aspects of the complete physics that are unknown), through numerical approximations (e.g., curve fits and response surfaces), or through variability within the physical processes.¹⁰⁻¹³ This aspect of uncertainty within the M&S effort is quite distinctive and unique, setting M&S apart from command, control, display, and support software.

While the principles of good software engineering generally carry over in the M&S community, the appropriateness of the model for an intended application, the uncertainties of the M&S effort, in general, and the credibility of a specific M&S product within a given application must also be assessed during development.¹⁻¹³ These additional concerns naturally lead the M&S community to consider aspects of uncertainty quantification (UQ) and uncertainty management (UM). The concepts of UQ and UM within M&S are a critical aspect of the decision-making processes. A companion paper in this session discusses how the rigor that is associated with several key aspects of an M&S effort (including the UQ and UM practices) influences the perceived credibility of that M&S in a given situation.¹⁴

Uncertainties in the technical performance metrics, along with those that are associated with schedule and budget, obviously introduce risk within any decision-making process. However, for decisions that are based upon M&S, the potential exists that the uncertainties (particularly of the technical performance metrics) may be significantly underestimated due to a wide range of factors; this potential for underestimated M&S uncertainty introduces additional risks that are distinct and that are typically not considered within NASA’s decision-making processes. The purpose of this paper is to propose a means to better illuminate the extent of the potential uncertainties that might be considered for decisions within an M&S effort.

In an M&S project, the software can execute correctly and do the job it was intended to do and yet still produce incorrect or invalidated results (i.e., results that do not correlate with the real world). Incorrect results can occur if the model that is used within the M&S software is incorrect because of approximations, unknown physics, numerical errors, or variability in the physical processes. That is, in the M&S community, the concept of software validation (i.e., appropriateness for the intended application) takes on a much larger meaning, which is generally unrecognized by the SE community. The M&S community has long struggled with the existing requirements, policy directives, procedures, guidelines, and best practices for software development because these documents contain no discussion of uncertainty; thus, these documents are largely ignored or paid “lip service” by most of the M&S community.

An M&S effort is often intended *to represent* an actual physical or imagined reality. Indeed, the Defense Modeling and Simulation Office (DMSO) Verification, Validation, and Accreditation Recommended Practices Guide (RPG), Build 3.0⁹, defines a simulation with the following language:

- i. "A technique for testing, analysis, or training in which...real-world and conceptual systems are reproduced by a model."
- ii. "An unobtrusive scientific method of inquiry involving experiments with a model rather than with the portion of reality this model represents."

Likewise, the RPG defines a representation as:

- i. "Something that stands in place of or is chosen to substitute for something else."

Recently, several small groups of M&S researchers and application practitioners from various organizations have attempted to address the issues of UQ and UM within decision-making processes. Within the U.S. Federal Government, significant efforts to address UQ and UM within the decision-making process have been undertaken by the Department of Energy (in particular at Sandia National Laboratory¹⁰), by the DMSO⁹ (now the Modeling and Simulation Coordination Office), and by NASA at Langley Research Center (LaRC)¹⁴⁻¹⁵ and Marshall Space Flight Center.¹⁶ A future effort will likely be made to compare and contrast the various ways these groups have chosen to approach the ideas of UQ and UM within decision-making processes. Perhaps some integration of the ideas and forms will be undertaken to produce one or more consolidated matrices.

This paper addresses M&S software implementations. These are typically used for technology investment decisions, concept evaluation, analysis, and design of flight systems. Some M&S are used during flight or mission operations to 1) monitor actual performance against predicted performance, 2) propose solutions to concerns that may arise, and 3) provide general support to the decision-making process. Another role of M&S is to support accident investigations following events of abnormal operations.

A simulation as used for an M&S project may exhibit varying degrees of correctness, depending upon a large and complex set of factors, including its intended use, the known physics and numerical approximations, and the referent data set against which the M&S correctness is compared; the latter may include experimental observations or flight data. The M&S has an associated element of *credibility*, which goes well beyond the service functions that are performed by command, controls, display, and support software. The credibility of the M&S has both objective and subjective parts. The objective part of M&S credibility is dependent on the specific disciplinary practices that are employed; these can be compared with accepted best practices as might be defined by a community of practice. The subjective part of M&S credibility is dependent upon the utility of specific M&S results to a decision maker in a given situation; this utility can really only be measured by a person who is charged with using the M&S results in a given situation to make a risk-informed decision.

To many people, the credibility of M&S results is closely related to the UM practices that are applied to the effort. Some of the key UM practices that are associated with the credibility of M&S include: verification, validation, uncertainty quantification, and error estimation. Because complex M&S software cannot be proven to be 100-percent correct for all inputs, these practices have degrees and attributes of correctness and applicability. The value and applicability of these practices and their attributes vary from circumstance to circumstance and from decision maker to decision maker.

This paper describes an uncertainty structure matrix for M&S projects, which enables one to obtain an objective description of the state of UM practices within a given M&S effort. The columns in the uncertainty structure matrix contain common UM practices, and the rows objectively describe levels of achievement for each practice. One can quickly look at the matrix to determine where a given M&S effort falls on a common set of UM practices, which are described in objective and absolute terms that can be applied to virtually any M&S effort. The matrix also can be used to plan those steps and resources that would be needed to advance one level of achievement in each category. The remainder of this paper is organized to describe the proposed uncertainty structure matrix and to provide some rationale for the choices made therein.

II. Description of the Proposed Uncertainty Structure Matrix

The uncertainty structure matrix that is proposed in this paper was developed by NASA LaRC researchers with somewhat diverse backgrounds: aerodynamics, radiation, and aerospace systems analysis/design. The uncertainty structure matrix development effort was completed within the larger context of developing a new (currently interim) NASA standard for models and simulations.¹⁵ This interim NASA standard is a part of the response that is required by the NASA Office of Chief Engineer (OCE) following the Columbia Accident Investigation Board (CAIB) report¹⁷ and the subsequent generalization to the NASA community by the Diaz Commission¹⁸ of the Shuttle-

specific return to flight recommendations, observations, and findings from the CAIB report. The interim standard and the proposed uncertainty structure matrix also respond to criticism that was voiced in the Return to Flight Task Group Final Report¹⁹ (see Annex A.2), which argued for broader usage of UQ and probabilistic methods within NASA.

Two key aspects of the interim NASA standard for models and simulations are requirements for 1) performing UQ activities for a given M&S effort, and 2) reporting to decision makers the extent and results of UQ activities performed for the M&S effort. The standard development team was seeking a means to objectively describe UQ and UM practices that could be broadly applied to a wide variety of M&S efforts, as might be found across NASA. The proposed uncertainty structure matrix has been presented to a variety of people outside the development team, and positive feedback has been received. The proposed uncertainty structure matrix is presented and described here, first in a generic form and then in a more detailed form. As shown in Table 1, the uncertainty structure matrix includes six columns that describe the canonical elements of UQ, which are generally common across many disciplines of M&S. The matrix also includes five rows, which describe progressively more rigorous levels of activity that can be undertaken to quantify and manage uncertainty within M&S applications. The intent, in all cases, was that the text in the cells of the matrix be inclusive of all disciplines of M&S projects, that is, someone working in any M&S discipline would be able to recognize aspects of their work in the matrix.

Ideally, M&S practitioners can identify where their own activities currently fall on the matrix and where they might like their activities to fall on the matrix at some point in the future (see examples shown in Table 1). Estimates potentially could be developed and presented to management and M&S developers in regard to the specific steps, cost, and schedule changes that would be required to move a particular M&S activity to a higher level of achievement (i.e., toward the bottom of the matrix) in each column. For the example shown in Table 1, the boxes labeled “M&S current state” reflect the developer’s and/or manager’s assessments of the current state of a single M&S effort, evaluated for each of the six canonical elements; boxes labeled “M&S future state” reflect the developer’s and/or manager’s assessments of a desired or planned future state of the same M&S effort. In some cases, as a result of programmatic priorities and resource limitations, the current state of an M&S project may be deemed acceptable; thus, no future improvements would be planned for that canonical element. The authors believe that a correlation exists between the degree of rigor in the UQ efforts that are performed and the degree of risk. Hence, for those M&S activities that are determined to be at a lower level of achievement (i.e., near the top of the matrix), the decision maker assumes more risk; for activities that are determined to be at a higher level of achievement (i.e., closer to the bottom of the matrix), the decision maker assumes less risk. Note, however, that the decision to attempt advancement of a given M&S project in a given canonical element should be strongly influenced by the cost/benefit ratio of the added steps that are required to accomplish the advancement of the activity to a higher level and compared with the benefit derived from such an advancement for the intended use; the allocation of resources within a given program structure must also be considered. That is to say that advancing a given M&S project one or more levels within any canonical element must be balanced with the other competing needs in the M&S effort and within the program and/or project funding for the effort.

As noted above, the matrix includes six columns, one for each of the following canonical elements:

Canonical elements:

- A. Code verification
- B. Parameter calibration
- C. Model validation
- D. Numerical error estimation
- E. Model error estimation
- F. Outcome adjustment

Code verification, shown as the first canonical element in the matrix, is similar to code verification as defined by the SE community, which includes progressive levels of rigorous testing and review to ensure that the software contains no coding errors. The next element, parameter calibration, is not generally considered by the SE community. Nearly every model that is developed or used within the M&S community contains one or more parameters (e.g., coefficients in equations, physical constants, or inputs that control the way the coded algorithm is executed) that must generally have values that are set from some external source. In the cases of equation coefficients and physical constants, these parameters take on values that tie the M&S software to the conditions that are required for the intended use, usually to model some physical process in the real world. However, if the parameters are not set to reflect the intended use, then the simulation will fail to produce credible results.

Model validation, the third element, is similar to validation as defined within the SE community. The process of validating the code ensures that it does what it is supposed to do, but it also includes an evaluation of the suitability and capability of the physical modeling as it is to be used for an intended application. This latter aspect of validation is not generally part of the SE validation process. This aspect of the validation is developed and demonstrated through a series of confidence-building exercises, which are conducted on unit problems and simplified system problems, and through investigations of certain aspects of the problem setup and implementation, including initial and boundary conditions, formulation of the modeled equations or behaviors, and the required geometric and physics fidelities that are needed for credible solutions for the intended purpose. Successful completion of each of these confidence-building exercises progressively advances the M&S project toward the capability to perform credible simulation of the physical processes for the intended purpose.

The processes that make up the remaining three canonical elements in the uncertainty structure matrix have no equivalent in the SE processes of verification or validation. Numerical error estimation involves the spatial, temporal, and, perhaps, statistical resolution of the physical phenomena that are used in M&S software, which significantly impacts the M&S credibility for a particular intended use. Model error estimation involves the assessment of how well the modeled physics of the M&S effort represent the intended physics; this is also known to some parts of the M&S community as the model form error and is generally difficult to assess. Finally, outcome adjustment has to do with moving a broadly verified and validated M&S effort that is developed for one intended use to another application domain through a series of one or more defensible changes. An example might be an aerodynamic database that has been developed for one configuration or flight condition that is later applied to a different configuration or flight condition through the use of defensible increments.

The first row of the uncertainty structure matrix (Table 2) provides a definition for each canonical element; these definitions are in the form of succinct statements that are broadly recognizable across the M&S community. The subsequent rows of the matrix present levels of increasing activity to decrease the risk assumed by the decision. The levels of activity range from essentially no process to an independent analysis conducted by a highly qualified Subject Matter Expert (SME) to verify and validate the conclusions of the original M&S practitioner.

General level definitions:

1. No process: verification and validation are clearly inadequate; uncertainty is unknown.
2. UM: confidence building process that is achieved through unit problems and comparisons. SME must be used.
3. UQ for validation: a defined, systematic, documented process is followed to bound and quantify uncertainty for validation.
4. UQ for prediction: a defined, systematic, documented process is followed to bound and quantify uncertainty for prediction. The structure of the problem is understood adequately for safe prediction.
5. Independent UQ for prediction: an additional, independent process is carried out for high-confidence prediction error quantification.

The initial level is the beginning point at which no work has been done to reduce risk or manage uncertainty. Level 2 uses a series of ad hoc confidence-building exercises and relies heavily upon SME input in the domain of the intended application to develop an argument that the M&S software is working correctly and credibly. Level 3 requires that a systematic and documented approach be used to demonstrate M&S software correctness and suitability for the intended application and to advance the argument that the M&S software is working correctly in a domain for which validation data are readily available and used for comparison. Level 4 requires that systematic and documented processes be used in combination with SME input to ensure that errors in a domain for which validation data are available and used for comparison are understood and correctly quantified; further, the physics in the prediction domain (i.e., M&S points for which no validation data is available) must be sufficiently understood, so that defensible M&S uncertainty predictions outside the domain of validation points can be readily made. Level 5 requires independent UQ for prediction; that is, the same conclusions that are reached by the M&S developers and practitioners are independently verified by one or more knowledgeable sources. This independent UQ must generally be conducted by a respected SME so that the subtleties and nuances of the verification, validation, and UQ that are performed by the M&S developers and practitioners are not lost or overturned by the independent practitioners. That is, the independent UQ agent should essentially be a potential consultant to the project, rather than an independent verification and validation (IV&V) person from the SE community, in order to be as credible as possible. However, if every aspect of the M&S project has been *sufficiently documented*, as would be required by the interim NASA standard for models and simulations, then an individual without specific knowledge of the M&S project or its intended application could potentially provide a credible IV&V assessment of the effort.

The steps that are described in the cells of the matrix represent a progressive set of practices that can be undertaken to quantify and manage uncertainty and risk within each of the canonical elements of the matrix. Although from an M&S credibility perspective, moving a specific M&S application toward the bottom of the matrix in each of the canonical elements is clearly desirable, this is often prohibitively expensive and time consuming. From a risk-management perspective, understanding the evidence that is presented about where a particular M&S application falls on the matrix for each canonical element and why the M&S effort is placed at that level of UQ and UM is often more important than the specific level at which the M&S effort falls for each of the canonical elements. Further, moving an M&S application toward the bottom of the matrix generally does not reduce the uncertainty of a given M&S result for an intended purpose, nor does it assure the M&S suitability for an application that is far removed from the validation data; rather, moving the M&S project toward the bottom of the matrix simply means that the processes that are used to obtain the uncertainty metrics of the M&S results employ techniques that can be objectively defined as being more credible by the M&S community.

Cell definitions:

- A. **Code verification:** The process of determining that a model implementation accurately represents the developer's conceptual description of the model and that numerically correct answers to the coded equations are obtained.
1. The code has not been verified for the intended uses.
 2. Verification has been performed for related problems, including unit problems and problems that are similar to the intended application. Comparisons have been made across codes for problems that are similar to the intended application.
 3. A systematic, documented process (including all relevant unit problem) has been carried out to verify the code for validation.
 4. A systematic, documented process (including all relevant unit problems) has been carried out to verify the code for prediction.
 5. An additional systematic, documented process (including all relevant unit problems) has been independently carried out to verify the code for the intended application.
- B. **Parameter calibration:** The process of optimization/adjustment of model parameters in the presence of numerical and experimental error.
1. The parameters have not been calibrated for the intended uses.
 2. Parameters have been calibrated for unit problems and are believed to be reasonable for the intended application.
 3. A systematic, documented process has been carried out to calibrate the model(s), including estimates of numerical and experimental error, for all relevant unit problems.
 4. Not applicable (no additional requirement beyond Level 3).
 5. An additional systematic, documented process has been independently carried out to verify the calibration of the code, including estimates of numerical and experimental error.
- C. **Model validation:** The process of determining the degree to which a model is an accurate representation of the real world.
1. The model has not been validated for the intended uses.
 2. Validation has been performed for related problems, including unit problems and problems that are similar to the intended application.
 3. A systematic, documented process has been carried out to validate the model(s), for all relevant unit problems, to include estimates of numerical and experimental error.
 4. A systematic, documented process has been carried out to validate the model structure for the intended application (including uncertainty and physics boundaries) adequately for safe prediction.
 5. An additional systematic, documented model validation has been independently performed by an SME.
- D. **Numerical error estimation:** The process of determining an estimate of the residual numerical error based on the computed results.
1. Numerical error estimation has not been performed for the intended uses.
 2. Convergence residuals have been estimated for problems that are similar to the intended application.

3. A systematic, documented process has been carried out to quantify the residual numerical error for the intended application.
4. Not applicable (no additional requirement beyond Level 3).
5. An additional systematic, documented process has been independently carried out to quantify the residual numerical error for the intended application.

As noted previously, numerical error estimation involves a strict examination of the modeled temporal, spatial, or statistical behaviors. Although this activity is most easily associated with M&S efforts in which partial differential equations are solved, the rigor of the practices that are described within this element can also readily encompass many M&S efforts that are not expressible in terms of partial differential equations. Still, this canonical element may not apply to all forms of M&S.

- E. **Model error estimation:** The process of estimating the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model.
 1. Model form error has not been estimated for the intended uses.
 2. Model form error has been estimated for unit problems. Comparisons have been made across models for problems that are similar to the intended application.
 3. A systematic, documented process has been carried out to quantify the model error based upon comparisons with available experimental data.
 4. A systematic, documented process has been carried out to quantify the prediction model error.
 5. An additional systematic, documented process has been independently carried out by an SME for the intended application to quantify the model error.

- F. **Outcome adjustment:** The process of adjustment/update of solutions based on either experimental data or higher-fidelity simulations.
 1. The adjustment uncertainty has not been estimated for the intended uses.
 2. The uncertainty of the adjustment process, and the uncertainty associated with the higher fidelity source that was used to provide input to the adjustment process, have been captured at least at level two across the other five canonical elements.
 3. A defined, systematic, documented process has been followed to bound and quantify uncertainty for the adjustment with respect to validation.
 4. A defined, systematic, documented process has been followed to bound and quantify uncertainty of the adjustment with respect to prediction. The structure of the problem is understood adequately for safe prediction of the adjustment.
 5. An additional, independent process has been carried out for high-confidence prediction error quantification for the adjustment.

III. Conclusions

A proposed uncertainty structure matrix has been presented and described. The matrix includes six columns, which describe canonical elements that are common across many disciplines of M&S. The canonical elements that are included in the matrix are: code verification, parameter calibration, model validation, numerical error estimation, model error estimation, and outcome adjustment. The matrix includes a definition of each of the canonical elements, and five rows that describe progressively more rigorous levels of activity that can be undertaken to quantify and manage uncertainty within M&S applications. For activities that are placed near the top of the matrix in a given column, the decision maker would assume more risk due to the potential that uncertainties have been underestimated because they have been estimated with less rigorous methods; for activities that can be placed closer to the bottom of the matrix in a given column, the decision maker would assume less risk because the uncertainties have been estimated with more rigorous methods. The intent, in all cases, was that text in the matrix be inclusive of all disciplines of M&S (i.e., practitioners working in any discipline of M&S would be able to recognize aspects of their work in the matrix). Furthermore, M&S practitioners can identify the points at which their own activities currently fall on the matrix and where they might like their activities to fall on the matrix at some point in the future; estimates can be presented to management or M&S developers regarding the specific steps, cost, and schedule that would required to move a particular activity toward the bottom of the matrix for given UM elements.

References

- ¹ *Defense System Software Development*, DOD-STD-2167A, February 1988.
- ² *Software Engineering Software Life Cycle Processes Maintenance - IEEE Computer Society*, IEEE 14764, March 2006.
- ³ *Quality Management and Quality Assurance Standards - Guidelines for the Application of ANSI/ISO/ASQC Q9001 to the Development, Supply, and Maintenance of Software*, ASQ Q9000-3, January 1991.
- ⁴ *Information Technology - Software Life Cycle Processes*, First Edition, Amendment 2, ISO 12207, November 2004.
- ⁵ *Software Assurance Standard*, NASA-STD-8739.8, July 2004.
- ⁶ *National Aeronautics and Space Administration Software Policy*, NPD 2820.1C, August 2005.
- ⁷ *National Aeronautics and Space Administration Software Engineering Requirements*, NPR 7150.2, September 2004.
- ⁸ *National Aeronautics and Space Administration Program and Project Management Processes and Requirements*, NPR 7120.5, March 2005.
- ⁹ *VV&A Recommended Practices Guide*, Build 3.0, Defense Modeling and Simulation Office, September 2006 (URL: <http://vva.dmsomil/>, accessed February 2008).
- ¹⁰ Pilch, M.: *Risk Informed Decisions: "Measuring and Communicating Progress in Predictive Capability"*, 2006 Heart Conference, Sunnyvale, California, March 6–10, 2006.
- ¹¹ Banks, J. (Ed.): *Handbook of Simulation*, New York: John Wiley & Sons, 1998.
- ¹² "Guide for Verification and Validation of Computational Fluid Dynamics Simulations," American Institute of Aeronautics and Astronautics, AIAA G-077, 1998.
- ¹³ "Guide for Verification and Validation in Computational Solid Mechanics," American Society of Mechanical Engineers, ASME V&V 10, New York, NY, 2006.
- ¹⁴ Blattnig, S. R.; Green, L. L.; Luckring, J. M.; Morrison, J. H.; Tripathi, R. K.; and Zang, T. A.: "Towards a Credibility Assessment of Models and Simulations," AIAA 2008-2156, April, 2008.
- ¹⁵ *National Aeronautics and Space Administration Standard for Models and Simulations*, NASA-STD-(I)-7009, December 1, 2006 (currently still in interim status).
- ¹⁶ Lin, J.; West, J. S.; Williams, R. W.; and Tucker, P. K.: "CFD Code Validation of Wall Heat Fluxes for a GO₂/GH₂ Single Element Combustor," 41st AIAA/ASME/SAE/ASEE Joint Propulsion Conference, Tucson, AZ, July 10–13, 2005.
- ¹⁷ Columbia Accident Investigation Board, *Columbia Accident Investigation Board Report*, Volume 1, U.S. Government Printing Office, August 2003 (URL: http://caib.nasa.gov/news/report/pdf/vol1/full/caib_report_volume1.pdf, accessed February 2008).
- ¹⁸ Diaz, A. V.; Hubbard, S.; Novak, V.; Asrar, G.; Earls, J.; Kennedy, J.: "A Renewed Commitment to Excellence: An Assessment of the NASA Agency-Wide Applicability of the Columbia Accident Investigation Board Report," National Aeronautics and Space Administration, Valador, Inc., February 9, 2004.
- ¹⁹ Return to Flight Task Group, *Return to Flight Task Group Final Report*, U. S. Government Printing Office, July 2005, (URL: http://returntoflight.org/assets/pdf/final_rftg_report/RTF_TG_Final_Report_Lo-Res.pdf, accessed February 2008).

Table 1. Generic Uncertainty Structure Matrix for an M&S

| | Canonical Elements | | | | | |
|----------------|----------------------|----------------------|------------------------------------|----------------------|----------------------|------------------------------------|
| | Element 1 | Element 2 | Element 3 | Element 4 | Element 5 | Element 6 |
| Level 1 | | | | M&S current state | | |
| Level 2 | M&S current state | | | | M&S current state | |
| Level 3 | | M&S current state | | M&S future state | M&S future state | |
| Level 4 | M&S future state | M&S future state | M&S current and future state | | | |
| Level 5 | | | | | | M&S current and future state |

Table 2. Proposed Uncertainty Structure Matrix with Element and Level Descriptions

| Canonical Elements | | | | | | |
|--|---|--|--|---|---|--|
| Code and Model Development | | | Prediction | | | |
| | Code Verification | Parameter Calibration | Model Validation | Numerical Error Estimation | Model Error Estimation | Outcome Adjustment |
| Ideal Time Order | First | Second | Third | Fourth | Fifth | Sixth |
| Definitions / Key Characteristics | The process of determining that a model implementation accurately represents the developer's conceptual description of the model and that numerically correct answers to the coded equations are obtained. | Parameter calibration: The process of optimization / adjustment of model parameters in the presence of numerical and experimental error. | The process of determining the degree to which a model is an accurate representation of the real world. | The process of determining an estimate of the residual numerical error based on the computed results. | The process of estimating the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model. | The process of adjustment / update of solutions based on either experimental data or higher-fidelity simulations. |
| Level 1 No process: verification and validation are clearly inadequate; uncertainty is unknown. | The code has not been verified for the intended uses. | The parameters have not been calibrated for the intended uses. | The model has not been validated for the intended uses. | Numerical error estimation has not been performed for the intended uses. | Model form error has not been estimated for the intended uses. | The adjustment uncertainty has not been estimated for the intended uses. |
| Level 2 Uncertainty Management: confidence building process that is achieved through unit problems and comparisons. SME must be used. | Verification has been performed for related problems, including unit problems and problems that are similar to the intended application. Comparisons have been made across codes for problems that are similar to the intended application. | Parameters have been calibrated for unit problems and are believed to be reasonable for the intended application. | Validation has been performed for related problems, including unit problems and problems that are similar to the intended application. | Convergence residuals have been estimated for problems that are similar to the intended application. | Model form error has been estimated for unit problems. Comparisons have been made across models for problems that are similar to the intended application. | The uncertainty of the adjustment process, and the uncertainty associated with the higher fidelity source that was used to provide input to the adjustment process, have been captured at least at level two across the other five canonical elements. |

| | | | | | | |
|---|--|--|--|---|--|---|
| <p>Level 3 Uncertainty Quantification for Validation: a defined, systematic, documented process is followed to bound and quantify uncertainty for validation.</p> | <p>A systematic, documented process (including all relevant unit problem) has been carried out to verify the code for validation.</p> | <p>A systematic, documented process has been carried out to calibrate the model(s), including estimates of numerical and experimental error, for all relevant unit problems.</p> | <p>A systematic, documented process has been carried out to validate the model(s), for all relevant unit problems, to include estimates of numerical and experimental error.</p> | <p>A systematic, documented process has been carried out to quantify the residual numerical error for the intended application.</p> | <p>A systematic, documented process has been carried out to quantify the model error based upon comparisons with available experimental data.</p> | <p>A defined, systematic, documented process has been followed to bound and quantify uncertainty for the adjustment with respect to validation.</p> |
| <p>Level 4 Uncertainty Quantification for Prediction: a defined, systematic, documented process is followed to bound and quantify uncertainty for prediction. The structure of the problem is understood adequately for safe prediction.</p> | <p>A systematic, documented process (including all relevant unit problems) has been carried out to verify the code for prediction.</p> | <p>Not applicable (no additional requirement beyond Level 3).</p> | <p>A systematic, documented process has been carried out to validate the model structure for the intended application (including uncertainty and physics boundaries) adequately for safe prediction.</p> | <p>Not applicable (no additional requirement beyond Level 3).</p> | <p>A systematic, documented process has been carried out to quantify the prediction model error.</p> | <p>A defined, systematic, documented process has been followed to bound and quantify uncertainty of the adjustment with respect to prediction. The structure of the problem is understood adequately for safe prediction of the adjustment.</p> |
| <p>Level 5 Independent Uncertainty Quantification for Prediction: an additional, independent process is carried out for high-confidence prediction error quantification.</p> | <p>An additional systematic, documented process (including all relevant unit problems) has been independently carried out to verify the code for the intended application.</p> | <p>An additional systematic, documented process has been independently carried out to verify the calibration of the code, including estimates of numerical and experimental error.</p> | <p>An additional systematic, documented model validation has been independently performed by an SME.</p> | <p>An additional systematic, documented process has been independently carried out to quantify the residual numerical error for the intended application.</p> | <p>An additional systematic, documented process has been independently carried out by an SME for the intended application to quantify the model error.</p> | <p>An additional, independent process has been carried out for high-confidence prediction error quantification for the adjustment.</p> |