NASA Standard for Models and Simulations: Philosophy and Requirements Overview

Steve R. Blattnig[*], James M. Luckring[†], Joseph H. Morrison[‡]
NASA Langley Research Center, Hampton, Virginia, 23681

Andre J. Sylvester[§]
NASA Johnson Space Center, Houston, Texas, 77058

Ram K. Tripathi[**] and Thomas A. Zang[††]
NASA Langley Research Center, Hampton, Virginia, 23681

Following the Columbia Accident Investigation Board report, the NASA Administrator chartered an executive team (known as the Diaz Team) to identify those CAIB report elements with NASA-wide applicability and to develop corrective measures to address each element. One such measure was the development of a standard for the development, documentation, and operation of models and simulations. This report describes the philosophy and requirements overview of the resulting NASA Standard for Models and Simulations.

Nomenclature

AIAA	=	American Institute of Aeronautics and Astronautics
ASME	=	American Society of Mechanical Engineers
CAIB	=	Columbia Accident Investigation Board
FAA	=	Federal Aviation Administration
NPD	=	NASA Policy Directive
NPR	=	NASA Procedural Requirement

I. Introduction

AS one of its many responses to the 2003 Space Shuttle *Columbia* accident,[1] NASA decided to develop a formal standard for models and simulations (M&S).[2] Work commenced in May 2005 and an interim version[3] was issued in late 2006. This interim version underwent considerable revision following an extensive Agency-wide review in 2007 followed by some additional revisions as a result of the review by the NASA Engineering Management Board (EMB) in the first half of 2008. The NASA Chief Engineer issued the resulting permanent version of the NASA M&S Standard[4] in July 2008.

Bertch, Zang, and Steele[5] provided a summary review of the development process of this standard up through the start of the review by the EMB. A thorough recount of the entire development process, major issues, key decisions, and all review processes is available in Zang et al.[6] The present paper is one of a pair providing a summary of the permanent version of the NASA M&S Standard. Its focus is the overall philosophy of the standard and an overview of the requirements. The companion paper[7] summarizes the Credibility Assessment Scale, a key feature of the standard.

---

[*] Physicist; Durability, Damage Tolerance and Reliability Branch; Mail Stop 188E.
[†] Senior Research Engineer; Configuration Aerodynamics Branch; Mail Stop 499; Associate Fellow AIAA.
[‡] Assistant Head; Computational AeroSciences Branch; Mail Stop 128; Senior Member AIAA.
[§] Chief; Integration Test Facilities Branch; Mail Stop EV3; Associate Fellow AIAA.
[**] Senior Research Scientist; Durability, Damage Tolerance and Reliability Branch; Mail Stop 188E; Associate Fellow AIAA.
[††] Chief Technologist; Systems Analysis & Concepts Directorate; Mail Stop 449; Associate Fellow AIAA.

This paper starts with general background on the NASA M&S Standard—the motivation for its development, its objectives, and its scope. This is followed by a review of the extant M&S guidance (in NASA and elsewhere) at the start of its development and a discussion of the fundamental choices that were made about its goal and vocabulary. Then the organization of the requirements is described and key requirements are discussed. Finally, recommendations for maturation of the NASA M&S Standard are presented. Verbatim quotes from the NASA M&S Standard and other documents are used liberally throughout this document to demonstrate the motivation and implementation of the standard.

II. Background of the NASA M&S Standard

A. Motivation

In the course of the *Columbia* accident investigation, the Columbia Accident Investigation Board (CAIB)[1] documented the use of one particular M&S code, known as Crater, during the STS-107 flight. This serves as a useful case study in that the example includes issues that are not unique to the Crater code but have broader implications to M&S usage. A brief synopsis follows.

The Crater model was originally developed during the Apollo program to assess impact damage (cratering) by micrometeoroids. It was modified between 1979 and 1985 to enable the analysis of impacts to the acreage tiles that cover the bottom of the shuttle. "When used within its validated limits, Crater provides conservative predictions (that is, Crater predictions are larger than actual damage). When used outside its validated limits, Crater's precision is unknown." (Ref. 1, p. 144)

Following the STS-107 launch, a foam impact analysis was initiated using Crater. The performing engineer had received formal training on Crater and was certified by the contractor but had only used the program twice before. In addition, the estimated size of the foam that was used for the analysis was 400 times larger than the validated input domain for Crater predictions. Crater predicted a penetration depth greater than the thickness of the tile. However, because the results of calibration tests with small projectiles showed that Crater predicted a deeper penetration than would actually occur and because the Crater model did not account for the increased density of the lower tile layer, the engineers judged that the actual damage from the piece of foam lost from STS-107 would not be as severe as Crater predicted and assumed that the debris did not penetrate the skin of the orbiter.

The Columbia Accident Investigation Board report[1] identified three findings that were a direct result of the Crater case study that are relevant to the development of the NASA M&S Standard:

> F6.3-10: The Team's assessment of possible tile damage was performed using an impact simulation that was well outside Crater's test database. The Boeing analyst was inexperienced in the use of Crater and the interpretation of its results. Engineers with extensive Thermal Protection System expertise at Huntington Beach were not actively involved in determining if the Crater results were properly interpreted.

> F6.3-11: Crater initially predicted tile damage deeper than the actual tile depth, but engineers used their judgment to conclude that damage would not penetrate the densified layer of tile. Similarly, RCC damage conclusions were based primarily on judgment and experience rather than analysis.

> F6.3-13: The assumptions (and their uncertainties) used in the analysis were never presented or discussed in full to either the Mission Evaluation Room or the Mission Management Team.

A fourth finding from the CAIB Report relevant to the development of the NASA M&S Standard was:

> F10.1-4: The FAA and U.S. space launch ranges have safety standards designed to ensure that the general public is exposed to less than a one-in-a-million chance of serious injury from the operation of space launch vehicles and unmanned aircraft.

The CAIB also made the following Shuttle-specific recommendation:

> R3.8-2: Develop, validate, and maintain physics-based computer models to evaluate Thermal Protection System damage from debris impacts. These tools should provide realistic and timely estimates of any impact damage from possible debris from any source that may ultimately impact the Orbiter. Establish impact damage thresholds that trigger responsive corrective action, such as on-orbit inspection and repair, when indicated.

The NASA Administrator chartered an executive team to identify the CAIB report elements with Agency-wide applicability and to develop corrective measures to address each element. This executive team became known as the Diaz Team. In its general discussion of these issues, the Diaz Team report[2] suggested the following:

> All programs should produce, maintain, and validate models to assess the state of their systems and components. These models should be continually updated and validated against experimental and operational data to determine appropriate courses of action and repair. The value of the models should be assessed with respect to their ability to support decision making in a timely way so as not to lead the decision maker to a conflict between costly action versus effective action in the interest of safety or mission success.
>
> Personnel need to be adequately trained in model use, limitations, and escalation procedures when issues arise. Engineers, when faced with results that defy "reality checks," should double check the model then raise their concerns.
>
> NASA policies recognize requirements for public safety. Those policies should be reviewed and the models used should be continually updated and assessed with respect to value in supporting timely decision making.

Soon after the release of the CAIB report, the NASA Administrator appointed the Return to Flight Task Group to provide an independent assessment of NASA's actions to implement the recommendations of the CAIB. In July 2005, the Return to Flight Task Group issued their report.[8] Annex A2 contained numerous concerns regarding the use of M&S. The following are some excerpts from that report (the bold face is not in the original but is added here to highlight important points):

> Standard engineering practice calls for objectives (**requirements and interface definitions**) to be established prior to development for any model or system of models, and processes and **criteria defined for validating and verifying the model's results**…Initially, we did not observe these normal processes being followed during the development of these models…
>
> The uncertainties in one model (or system) inherently feeds into and compounds the uncertainty in the second model (or system), and so on. It appears, however, that NASA largely designed these five classes of models without the attention to the interdependencies between the models necessary for a complete understanding of the end-to-end result. Understanding the characteristics of, and validating and verifying, one type of model without examining the implications for the end-to-end result is not sufficient…But, as the Columbia accident showed, in a high risk environment that involves many unknowns like human space flight, **experience and instinct are poor substitutes for careful analysis of uncertainty**.
>
> …during the return-to-flight effort, there has been an enormous expenditure of time and resources—amounting to tens of millions of dollars—without the discipline of a formal **development plan**, clear objectives, **explicit plans for verification and validation**, thorough **outside review**, documented ICDs [interface control documents] between models, or a good understanding of the limitations of analytical systems employing **multiple, linked deterministic models**. Validation and verification planning has been left to the end of the process rather than the beginning…Analytical models have essentially driven the return-to-flight effort; however, industry and academic **standards and methods for developing, verifying, and validating the models have not been used**. In addition, no sensitivity analyses had been conducted and no empirical data from flight history had been incorporated in the models or their validation.

All three groups—the CAIB, the Return to Flight Task Group, and the Diaz Team—identified a need for NASA to establish greater discipline in its development and usage of models and simulations.

B. Objectives for the NASA M&S Standard

The Diaz Team, in Action #4 from its January 30, 2004 report,[2] explicitly called for NASA to: "Develop a standard for the development, documentation, and operation of models and simulations." Six specific objectives were associated with Action #4. These objectives are as follows:

1. Identify best practices to ensure that knowledge of operations is captured in the user interfaces (e.g. users are not able to enter parameters that are out of bounds).
2. Develop process for tool verification and validation, certification, re-verification, revalidation, and recertification based on operational data and trending.
3. Develop standard for documentation, configuration management, and quality assurance.
4. Identify any training or certification requirements to ensure proper operational capabilities.
5. Provide a plan for tool management, maintenance, and obsolescence consistent with modeling/simulation environments and the aging or changing of the modeled platform or system.

American Institute of Aeronautics and Astronautics

6. Develop a process for user feedback when results appear unrealistic or defy explanation.

Subsequently, the NASA Chief Engineer (then Christopher Scolese) augmented this in an internal memo dated Sept. 1, 2006, with the additional objective that "the M&S standard will

7. Include a standard method to assess the credibility of the M&S presented to the decision maker when making critical decisions (i.e., decisions that effect human safety or mission success) using results from M&S."

The Chief Engineer's expectation was that the "M&S standard will …establish M&S requirements and recommendations that will form a strong foundation for disciplined (structure, management, control) development, validation and use of M&S within NASA and its contractor community."

As stated in Ref. 4 these seven "objectives are encapsulated in the overall goal for the M&S Standard, which is to ensure that the credibility of the results from M&S is properly conveyed to those making critical decisions. Critical decisions based on M&S results, as defined by this standard, are those technical decisions related to design, development, manufacturing, ground, or flight operations that may impact human safety or program/project-defined mission success criteria."

C. Scope of the NASA M&S Standard

The determination of those M&S that fall within the scope of the NASA M&S Standard is based upon an assessment of the risk that is posed by the use of the M&S. Fig. 1, taken from appendix A of the NASA M&S Standard, illustrates the approach. Those M&S that are high consequence and high influence are shown in the red boxes in Fig. 1 and are within the scope of the NASA M&S Standard; those that are low consequence or low influence are shown in the green boxes and are outside the scope; and those that have intermediate consequence and intermediate influence are shown in the yellow boxes and may be categorized at the discretion of the program and technical authorities. Ref. 9 provides the context for the phrase "technical authority" that appears here and in several requirements in the standard:

> NASA's success is dependent upon a proper balance between those authorities vested in program and project managers intended to promote programmatic efficiency, and those authorities vested in institutional managers intended to assure resource availability, compliance with external requirements, compliance with applicable standards of professional practice, and efficiency across NASA's total program portfolio...[the institution] provide[s] individuals who have a formally delegated Technical Authority role traceable to the Administrator and are funded independent of Programmatic Authority. The Technical Authorities are a key part of NASA's overall system of checks and balances and provide independent oversight of programs and projects in support of safety and mission success.

This relationship is demonstrated in Fig. 2.

M&S Results Influence		IV: Negligible	III: Marginal	II: Critical	I: Catastrophic
	5: Controlling	(G)	(Y)	(R)	(R)
	4: Significant	(G)	(Y)	(R)	(R)
	3: Moderate	(G)	(Y)	(Y)	(R)
	2: Minor	(G)	(G)	(Y)	(Y)
	1: Negligible	(G)	(G)	(G)	(G)
		IV: Negligible	III: Marginal	II: Critical	I: Catastrophic
		Decision Consequence			

Figure 1. Sample M&S risk assessment matrix.

The NASA M&S Standard is an institutional standard under the auspices of the NASA Office of the Chief Engineer; hence, the relevant technical authority is the engineering authority. In general, the (engineering) technical authority is responsible for ensuring that the NASA M&S Standard is employed properly on those NASA programs for which it is applicable. The determination of which M&S fall within scope of the standard using the M&S Risk Assessment is but one example of a responsibility that is shared by the program and the technical authority.

Each of the NASA mission directorates has the programmatic responsibility for several programs. A program typically consists of several projects; in turn, modeling and simulation is but one of the many elements of a project. Programmatic decisions related to M&S are made at various levels in this hierarchy. In this paper, we usually use the term "project" in the generic sense, referring to the appropriate level in the programmatic hierarchy.
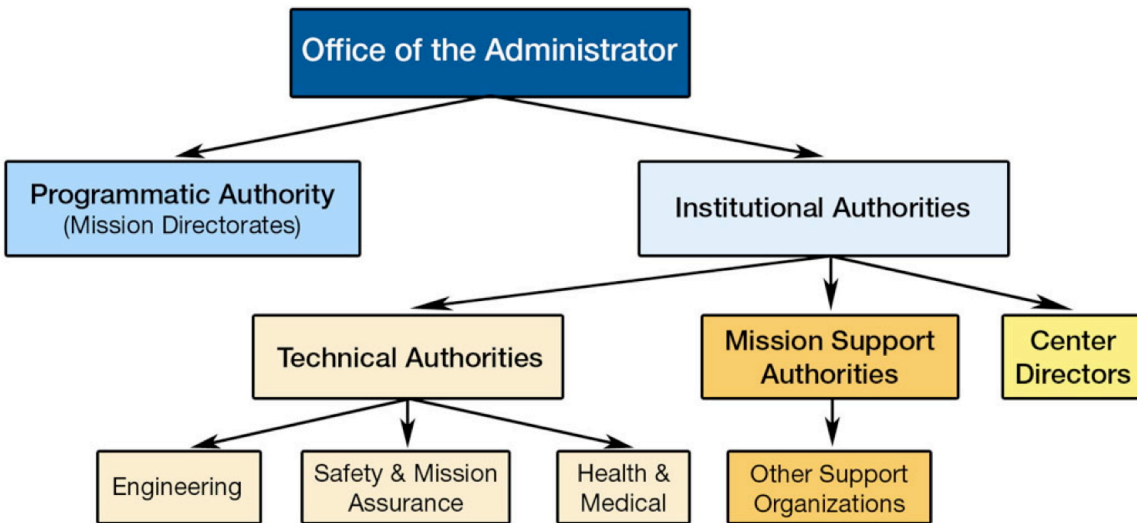
Figure 2. Relationship between NASA programmatic and institutional authorities.

III. Fundamental Approach to the NASA M&S Standard

A. Review of Existing M&S Guidance

A thorough review of available guidance for M&S both inside and outside NASA at the inception of the response to Diaz Action #4 revealed that although considerable guidance is available for the development, operations (use), and management of M&S, no formal, requirements-focused standard exists. Examples of such guidance include the Defense Modeling and Simulation Office (DMSO) VV&A Recommended Practices Guide,[10] the American Institute of Aeronautics and Astronautics (AIAA) Guide for Verification and Validation of Computational Fluid Dynamics Simulation,[11] and the American Society of Mechanical Engineers (ASME) Guide for Verification and Validation in Computational Solid Mechanics.[12] None of these contain actual requirements, nor do they deal to any significant degree with the operation and management of M&S but rather focus on development of M&S.

The following findings (documented in Zang et al.[6]) resulted from this review:

F-1. Current NASA standards are strongly oriented towards control systems and displays. Quality assurance and configuration management are very well covered, but the unique, critical aspects of models and simulations are not addressed, for example, validation against experimental or flight data, and uncertainty quantification.

F-2. No federal agency has an M&S Standard, although the DoD has extensive M&S guidance, and the Nuclear Regulatory Commission has standards for control systems and displays.

F-3. Relevant M&S guidance is strongly focused on the development phase of the M&S life-cycle, and especially upon verification and validation. There is little guidance on the operations of M&S and virtually no guidance on the maintenance of M&S.

F-4. NASA has no policy, nor any procedural requirements for M&S, except for the software engineering aspects of M&S covered by NPD 2820.1B[13] and NPR 7150.2.[14]

B. Basic Philosophy

A major challenge for the formulation of the NASA M&S Standard was devising an approach to requirements that was suitable for all types of M&S and a broad range of applications. This challenge was met by the philosophical choice that the standard should focus on the requirements for the documentation and reporting of M&S results to decision makers. For the most part, the standard specifies *what* shall or should be done; it does not prescribe *how* the requirements are to be met. (The one exception is the use of the Credibility Assessment Scale, which is described in the companion paper.[7]) Because the standard is applicable across widely differing disciplines

and applications and has a wide scope, it could not be prescriptive. A single high-level document like a standard could not contain all of the necessary details to be prescriptive while still covering all of the applications that are relevant to the scope. For example, solution verification is accomplished differently for a computational fluid dynamics problem than it is for a radiation analysis problem. The decision was made that technical details were better left to the appropriate technical authorities and the recommended practices guides. The focus of the standard was, thus, not the determination of how these disciplines should be performing their work but on how to communicate what was done so that decision makers have the necessary information on which to base their decisions.

Another distinct choice that was made was to employ the concepts and language of the modeling and simulation community rather than those of the software engineering or the systems engineering communities. Indeed, many of the concepts and much of the language of modeling and simulation are related to the scientific method, which traces back nearly a millennium, for example, to Ibn al-Haytham[15] (1021), Roger Grosseteste[16] (1235), and Roger Bacon[17] (1267). Some of the key aspects of the scientific method that are captured in the M&S activities are characterization and hypotheses (modeling), predictions (simulation), and experiments (validation). Software engineering (circa 1980) and systems engineering (circa 1960) are much younger antecedents to M&S. More importantly, the target audience for the NASA M&S Standard consists of M&S practitioners; the best means of communicating effectively with them is using their language. This choice manifests itself most clearly in the definitions of verification and validation. Table 1 compares the definitions of verification and validation in official NASA documents for systems engineering (NPR 7123.1[18]), software engineering (NPR 7150.2[14]), and M&S (NASA-STD-7009[4]). The latter definitions are minor modifications of the definitions in the ASME Guide,[12] which, in turn, evolved from those in the AIAA Guide[11] and the U.S. Department of Defense (DoD) Recommended Practices Guide (RPG).[10] Of particular note is that the software engineering community's definitions of verification and validation which emerged in the 1980's,[19] were deemed inappropriate to M&S by the DMSO, the AIAA, and the ASME. See Oberkampf and Trucano[20] for a thorough discussion of the history of the definitions of verification and validation for M&S.

Table 1. Definition of verification and validation.

Document	Verification Definition	Validation Definition
NPR 7123.1A (reflects systems engineering perspective)	Proof of compliance with specifications. Verification may be determined by test, analysis, demonstration, and inspection.	Proof that the product accomplishes the intended purpose. Validation may be determined by a combination of test, analysis, and demonstration.
NPR 7150.2 (reflects software engineering perspective)	Software verification is a software engineering activity that demonstrates that the software products meet specified requirements.	Software validation is a software engineering activity that demonstrates that the as-built software product or software product component satisfies its intended use in its intended environment.
NASA-STD-7009 (reflects M&S perspective)	The process of determining that a computational model accurately represents the underlying mathematical model and its solution from the perspective of the intended uses of M&S.	The process of determining the degree to which a model or a simulation is an accurate representation of the real world from the perspective of the intended uses of the model or the simulation.

These definitions are context dependent and reflect accepted practices within each community. For both the systems engineering and the software engineering communities, verification entails testing against product specifications, and validation entails determining whether the product is sufficient for the intended use. Both of these result in a "yes" or "no" decision, i.e., either the product (systems engineering) or software (software engineering) meets specifications or it does not. For the M&S community, the verification process includes testing to determine how well a computational model represents the underlying mathematical model and the validation process includes testing to determine how well the model represents the real world. M&S verification and validation is an ongoing process; it generally does not result in a "yes" or "no" decision.

The U.S. DoD, which is the largest sponsor and user of M&S in the world,[21] has issued a policy governing M&S for DoD usage. This policy covers the verification, validation, and accreditation of models and simulations; it emphasizes the accreditation of models and simulations for specific purposes.[22] Balci[21] has developed a methodology to certify the acceptability of modeling and simulation based on a body of methods, rules and postulates for employment of subject matter experts; constructing a hierarchy of indicators; relative criticality weighting of indicators; using a rule based expert knowledge software tool; assigning crisp, fuzzy, and nominal scores for indicators; aggregating indicator scores; graphically representing the scores and weights; preparing the certification report; and interpreting the results. This approach specifies how the requirements are to be met throughout the life cycle of M&S and is designed to measure the acceptability of the M&S to the decision maker. The acceptability of the M&S includes measures of credibility, risk, and cost. The NASA M&S Standard is focused on the credibility of specific M&S results rather than the credibility of a class of M&S. The risk of the use of the M&S results is captured in the M&S Risk Assessment Matrix. The cost of the M&S is not addressed in the NASA M&S Standard, as it is germane to neither the credibility of the results nor the risk that is associated with their use. The NASA M&S Standard did not address *how* to meet the requirements because the field of M&S, especially verification, validation, and uncertainty quantification, has not reached maturity, and new and better approaches could be developed as the field matures. Balci[21] also includes attributes as metrics that don't directly affect credibility, such as efficiency, interoperability, and maintainability, and requires an assessment of the process that is used to create the product throughout its life cycle; this feature would not be possible with commercial, off-the-shelf software for M&S.

Critical decisions that are based entirely or partially on M&S are usually made within the context of a project and are often based on a combination of factors including several M&S results. The decision maker, having an overarching view of the project, has the responsibility for accepting the M&S results with full awareness of their impact on the whole project. Credible results play a crucial role in the process. To facilitate the decision, the practitioners report the results in a transparent manner, thus ensuring the repeatability, robustness, and traceability of the reported results. The magnitudes and confidence levels of the uncertainties are needed to make risk assessments and to make compromises, if need be, in the decision.

More details on the fundamental approach to the NASA M&S Standard are given in Sect. 7.2.4. of Zang et al.[6]

IV. Overview of Requirements

The requirements section consists of 49 requirements, which are organized into eight subsections. The first seven subsections (4.1–4.7) provide the underlying activities for the eighth subsection (4.8), "Reporting of M&S results to decision makers," which is the main emphasis of the standard. The first seven subsections set requirements mainly to establish the type of information that needs to be collected so that it can be effectively communicated to those making decisions. A complete list of the requirements and their traceability to the seven objectives given in section II of this paper can be found in Zang et al.[6] This section summarizes the main points.

Thirty-three of the requirements start with the words "shall document." Twelve of these (i.e., 4.1.5, 4.2.6, 4.2.8, 4.3.6, 4.4.1, 4.4.2, 4.4.4–4.4.9) are to be interpreted as meaning that the activity in question is not required per se, but that whatever was done is to be documented, and if nothing was done a clear statement to that effect is to be documented.

The first requirements subsection (4.1) addresses programmatic activities. The most fundamental activity is for project management, in collaboration with the technical authority, to identify and document the critical decisions that are to be addressed with M&S and determine which M&S are within the scope of the NASA M&S Standard. The latter determination should be based upon the risk that is posed by the anticipated use of the M&S (see Fig. 1 above and the related discussion). These requirements oblige the project to 1) identify the M&S that are in scope, 2) define the objectives and requirements for the M&S, and 3) develop a plan for the acquisition, development, operation, maintenance, and/or retirement of the M&S.

The second requirements subsection (4.2) contains the requirements that are associated with the model, where model refers to the conceptual model, the mathematical model, and the computational model. The majority of these requirements address documentation for the assumptions; the basic structure; mathematics; data sets; limits of operation; guidance in the proper use of the model; parameter calibrations; model updates; and methods for the uncertainty quantification for any data that are used to develop the model or that are incorporated into the model.

The third requirements subsection (4.3) contains the requirements that are imposed on the simulation. The NASA M&S Standard states: "The execution (operation) of a computational model and the processing of the results from the simulation are simulation and analysis activities, respectively." This includes those requirements that address the limits of operation, the pedigree of the input data, the assessment of the appropriateness of the

American Institute of Aeronautics and Astronautics

simulation relative to its intended use, and the use history of the M&S, as well as processes for executing the simulations and for conducting analyses.

The fourth requirements subsection (4.4) addresses verification, validation, and uncertainty quantification. M&S practitioners typically understand the nuances of these requirements for their particular type of M&S. Specific emphasis is given to communicating the domains of verification and validation of the model to ensure appropriate application of the model. Furthermore, documentation of the uncertainties in the results and the sensitivities is required. All but one of the requirements in this subsection are descriptive rather than prescriptive.

The fifth requirements subsection (4.5) of the standard addresses the use of recommended practices. The sixth requirements subsection (4.6) addresses training for developers, operators, and analysts. (Both topics were explicitly specified in the Diaz Action #4.)

The seventh requirements subsection (4.7) addresses the Credibility Assessment Scale. The requirements specify that the M&S results and processes be assessed on the Credibility Assessment Scale, the details of which are defined in appendix B of the standard. (See Ref. 7.)

The eighth and final requirements subsection (4.8) addresses the reporting of results to decision makers. This is the focal point of the NASA M&S Standard.

V. Discussion of Some Key Requirements

Many of the requirements in the NASA M&S Standard are nothing more than M&S-specific instances of good engineering practice. In this section we highlight those requirements that call for special mention. All requirements listed below are taken verbatim from the NASA M&S Standard.[4]

> **Req. 4.1.1**–Shall document the risk assessment for any M&S used in critical decisions.

Section II of this paper discusses the M&S risk assessment that forms the basis for determining whether a particular M&S falls within the scope of the NASA M&S Standard. The onus for performing this risk assessment rests with project management (with the participation of the technical authority). The M&S risk assessment matrix shown in Fig. 1 is patterned after the overall risk assessment approach that is used for all NASA projects. Appendix A of this paper provides detailed definitions of the decision consequence and M&S results influence axes. The former are minor adaptations of the consequence definitions given in NPR 8000.4,[23] whereas the latter were developed specifically for the NASA M&S Standard. In summary, the more serious the consequences of the decision and the more the decision is influenced by M&S results, the greater the risk that is associated with the use of the M&S results.

> **Req. 4.1.3**–Shall define the objectives and requirements for M&S products including the following: (a) The acceptance criteria for M&S products, including any endorsement for the M&S. (b) The rationale for the weights used for the subfactors in the Credibility Assessment Scale. (c) Intended use. (d) Metrics (programmatic and technical). (e) Verification, validation, and uncertainty quantification. (f) Reporting of M&S information for critical decisions (see section 4.8). (g) CM [Configuration Management] (artifacts, timeframe, processes) of M&S.

Despite the language in objective 2 of Diaz Action #4, the NASA M&S Standard does not mandate certification (called "accreditation" in some circles) of M&S. The decision on whether any type of "endorsement," such as but not limited to certification, is required for M&S within scope of the standard is left to the project (and the technical authority). This approach was taken at the direction of the Office of the Chief Engineer. The rationale is that formal endorsement is not always appropriate. The technical authority has a voice in determining these objectives and requirements; the preamble to the requirements in sect. 4.1 states: "the Technical Authority has particular responsibility to assure appropriate outcomes of Req. 4.1.3." The requirement impacts all categories and plays a pivotal role in the Credibility Assessment Scale to assure that M&S results are aligned for the intended use.

> **Req. 4.1.7**–Shall document the extent to which an M&S effort exhibits the characteristics of work product management, process definition, process measurement, process control, process change, and continuous improvement, including CM [Configuration Management] and M&S support and maintenance.

This requirement has ties to both objective 5 of Diaz Action #4 and to the M&S management factor in the Credibility Assessment Scale. Note that it covers the entire life cycle of the M&S, which is a distinguishing aspect of this standard; in particular, the NASA M&S Standard covers not just the development of the M&S but also the operation and maintenance. In contrast, most guidance on verification and validation concentrates almost exclusively on the development phase of the M&S.

Req. 4.2.5–Shall document the limits of operation of the model.

The term "limits of operation" is defined as "the boundary of the set of parameters for which an M&S result is acceptable based on the program/project-required outcomes of verification, validation, and uncertainty quantification." This requirement is driven by objective 1 of Diaz Action #4. Its purpose is to ensure that documentation is available so that those groups that are developing models and conducting simulations can take reasonable precautions to assure that simulations are not conducted with parameters that are out of bounds for the models. The following requirement ensures that the decision maker is informed if no guarantee that the simulation was indeed performed within the limits of operation of the model is available.

Req. 4.3.1–Shall do either of the following:
a. Ensure that simulations are conducted within the limits of operation of the model, or
b. Placard the simulation and analysis results with a warning that the simulation may have been conducted outside the limits of operation and include the type of limit that may have been exceeded, the extent that the limit might have been exceeded, and an assessment of the consequences of this action on the M&S results.

Requirements 4.2.5 and 4.3.1 tie to findings F6.3-10, F6.3-11, and F6.3-13 of the CAIB report and to objective 1 of Diaz Action #4; both of these findings and the associated Diaz action are a result of the Crater case study that was previously described in section 2.

Req. 4.2.7– Shall document guidance on proper use of the model.

"Guidance on proper use of a model includes descriptions of appropriate practices for set-up, execution, and analysis of results." Often models are complicated enough that the correct use is not obvious. For example, codes may only work properly on certain platforms with certain compilers; model options may need to be tailored specifically to the application of interests; or output from codes may need to be post-processed in order to be analyzed properly. Mistakes at any point in the execution of the model or the analysis of the output can produce invalid results. To help prevent these types of errors, the standard processes from model setup, execution, and analysis of results need to be documented.

Req. 4.5.1– Shall identify and document any Recommended Practices that apply to M&S for the program/project.

The standard was written to address issues that are broadly applicable to M&S and, as such, implementation details that are specific to particular disciplines, codes, or projects were purposely excluded. These details are typically embodied in recommended practices guides. Recommended practices are generally expected to evolve faster than overarching governance from the standard, as these practices are more closely coupled to M&S technological advances. Requirement 4.5.1 is designed to assure that applicable recommended practices are identified and documented.

Req. 4.8.1–Reports to decision makers shall include explicit warnings for any of the following occurrences, accompanied by at least a qualitative estimate of the impact of the occurrence:
a. Unfavorable outcomes from or failure to perform the intended use and setup/execution assessments (described in Req. 4.3.9 and Req. 4.3.10).
b. Any unachieved acceptance criteria (as specified in Req. 4.1.2).
c. Waivers to any of the requirements in this standard.
d. Violation of any assumptions of any model.
e. Violation of the limits of operation.
f. Execution warning and error messages (see Req. 4.3.2).

Req. 4.8.2–Reports to decision makers of M&S results shall include an estimate of their uncertainty and a description of any processes used to obtain this estimate as defined in Req. 4.4.9 and Req. 4.4.10.
a. 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.

Req. 4.8.3–Reports to decision makers shall include the level of credibility for the M&S results and the subfactor weights, using the process specified in section 4.7.

American Institute of Aeronautics and Astronautics

As stated previously, the main goal of the standard is to ensure that the decision maker is made aware of the key information regarding M&S results that is needed to infer their credibility. The information needed was broken down into three parts: the uncertainty of the results, the assessment on the scale, and any caveats that go along with the results. Requirement 4.8.2 prescribes that an explicit statement be made on the uncertainty in the results. However, this requirement leaves the presenter with the option of stating that no uncertainty estimate is available. This option is provided because circumstances (time pressure during mission operations or absence of uncertainty estimation capability) may preclude such estimates. Requirements 4.8.1 and 4.8.3 ensure that the decision maker has the information that is needed to determine the trustworthiness of the results including the uncertainty estimation. At this point, the decision maker must determine how much weight to give the M&S results.

VI. Recommendations for Maturation

As the NASA M&S Standard was being completed, its authors recognized that follow-on activities to the standard development would enhance the usefulness of the standard to agency M&S development and operations. Twelve recommendations for maturing the standard and the body of knowledge for effective models and simulation development were proposed by Zang et al.,[6] six of which are reprinted below:

R-1. NASA should integrate the M&S Standard into the NASA guidance hierarchy. The initial review of existing M&S guidance and standards (Section 7.2.1 [of Ref. 6]) made it apparent that the M&S Standard was not tied to any existing NPD [NASA Policy Directive] or NPR [NASA Procedural Requirement]. The most logical existing NPR that could link to the M&S Standard would be NPR 7123.1A (NASA Systems Engineering Processes and Requirements). Either the M&S Standard should be linked to an NPD, a future version of NPR 7123.1A or to a forthcoming NPR on NASA standards.

R-2. NASA should coordinate with other organizations and professional societies to further mature the M&S Standard. The development and operation of M&S, the analysis and presentation of M&S results, the proper training of M&S practitioners, the identification of recommended practices, and the need for assessing and conveying the credibility of M&S results to decision makers are not unique to NASA. These aspects of the M&S process are common to many other organizations. NASA should participate in activities directed towards standards that serve a broader M&S community.

R-3. NASA should sponsor development of Recommended Practices Guides. While some M&S have well established and documented procedures, many others do not. Furthermore, existing guidelines may not cover new applications of the M&S. For example, models often require calibration, or numerical parameters need to be tuned for new problems. Knowledge of these procedures, calibrations, and tunings often resides in a small subset of workers. NASA should identify M&S domains that need Recommended Practice Guides and coordinate with professional societies, academia, commercial and international partners to develop them. (Domains may be organized according to type of M&S, by discipline, or by application.)

R-6. Information regarding credibility assessment scale usage should be collected to determine effectiveness and provide data for further revision. In general, scales measuring the rigor, credibility, or similar aspects of M&S results have not received much use, and there is no consensus on such assessments. In particular, the credibility assessment scale in the M&S Standard has not been used. The immaturity of this particular field necessitates close monitoring of the impact of credibility assessment scale usage by NASA programs and the use of that information to update the credibility assessment scale. This is not a criticism of the present credibility assessment scale, but merely an acknowledgment of the state of such assessments; operational use is essential to advance the state–of–the-art.

R-9. NASA should collect data on the scope decisions, the cost impact and the credibility assessment scale usage of the M&S Standard. This is a more general recommendation than R-6. The extension to collection of data on the scope decisions and cost impact was motivated by the large number of comments on these topics submitted as part of the Agency-wide review.

R-10. NASA should develop, by application domain, an M&S "validation lessons learned" database. This information would be used to develop guidelines allowing designers to intelligently balance risk versus conservatism during program/project formulation. Solid data and rationale for design margins exist, in the form of written guidelines at the agency level, for only a few of the many application domains (i.e., disciplines). Of particular interest is the knowledge of why and by how much M&S results were in error before the models were tuned/correlated. NASA should also implement a process by which the guidelines are continuously re-evaluated and updated as the database grows.

VII. Concluding Remarks

NASA developed the M&S Standard in response to the CAIB findings with respect to models and simulations that are used for decision making. The overall philosophy of the standard is to document those artifacts of M&S products and tools that aid in determining the credibility of the M&S results. This documentation provides the basis for decision makers to assess the credibility of the M&S and to then apply the M&S results for critical decision making. With the development of the Credibility Assessment Scale, the standard provides a framework for evaluation of M&S products.

Appendix

The following subsections are taken verbatim from the NASA M&S Standard.[4]

A.1 Decision Consequence
Consequence classifications assess the impact of a decision that proves incorrect. The number of Consequence levels and most of the language is taken from NPR 8000.4. The last item in each class description has been added to address impact upon mission success criteria, such as science objectives.

   a. **Class IV - Negligible**. A poor decision may result in the need for minor first aid treatment but would not adversely affect personal safety or health; damage to facilities, equipment, or flight hardware more than normal wear and tear level; internal schedule slip that does not impact internal development milestones; cost overrun less than 2 percent of planned cost; all mission success criteria met, with at worst minor performance degradations.

   b. **Class III - Moderate**. A poor decision may result in minor injury or occupational illness, or minor property damage to facilities, systems, equipment, or flight hardware; internal schedule slip that does not impact launch date; cost overrun between 2 percent and not exceeding 15 percent of planned cost; a few (up to 25 percent) mission success criteria not met due to performance degradations.

   c. **Class II - Critical**. A poor decision may result in severe injury or occupational illness, or major property damage to facilities, systems, equipment, or flight hardware; schedule slippage causing launch date to be missed; cost overrun between 15 percent and not exceeding 50 percent of planned; many (between 25 percent and 75 percent) mission success criteria not met due to substantial performance degradations.

   d. **Class I - Catastrophic**. A poor decision may result in death or permanently disabling injury, facility destruction on the ground, or loss of crew, major systems, or vehicle during the mission; schedule slippage causing launch window to be missed; cost overrun greater than 50 percent of planned cost; most (more than 75 percent) mission success criteria not met due to severe performance degradations.

A.2 M&S Influence
Influence estimates the degree to which M&S results influence program/project engineering decisions. (Engineering decisions include determination of whether design requirements have been verified.)

   a. **Influence 1 - Negligible**. Results from the M&S are a negligible factor in engineering decisions. This includes research on M&S methods, and M&S used in research projects that have no direct bearing on program/project decisions (for NASA missions).

   b. **Influence 2 - Minor**. M&S results are only a minor factor in any program/project decisions. Ample flight or test data for the real system in the real environment are available, and M&S results are used just as supplementary information.

   c. **Influence 3 - Moderate**. M&S results are at most a moderate factor in any program/project decisions. Limited flight or test data for the real system in the real environment are available, but ample flight or test data for similar systems in similar environments are available.

   d. **Influence 4 - Significant**. M&S results are a significant factor in some program/project decisions, but not the sole factor for any program/project decisions. Ample flight or test data for similar systems in similar environments are available.

   e. **Influence 5 - Controlling**. M&S results are the controlling factor in some program/project decisions. Neither flight nor test data are available for essential aspects of the system and/or the environment.

Acknowledgments

References

[1]Gehman, H. W., et al., "Columbia Accident Investigation Board Report, Volume 1," August 2003, URL: http://caib.nasa.gov/.

[2]Diaz, A., et al., "A Renewed Commitment to Excellence. An Assessment of the NASA Agency-wide Applicability of the Columbia Accident Investigation Board Report," January 30, 2004 (P. A-5, R13) (http://www.nasa.gov/pdf/55691main_Diaz_020204.pdf).

[3]NASA, "Standard for Models and Simulations," NASA-STD-(I)-7009, December 1, 2006, URL: ftp://ftp.sonic.net/pub/users/usacm/PTC60/NASA_STD_I_7009.pdf.

[4]NASA, "NASA Standard for Models and Simulations," NASA-STD-7009, July 11, 2008, URL: http://standards.nasa.gov/public/public_count.taf?system=PUBLIC&event=PUBLIC_DD&Linkid=6365.

[5]Bertch, W. J., Zang, T. A., and Steele, M. J., "Development of NASA's Models and Simulations Standard," Paper 08S-SIW-037, 2008 SISO Spring Simulation Interoperability Workshop, Providence, RI, April 13–18, 2008.

[6]Zang, T. A., et al., "NASA Standard for Models and Simulations (M&S): Development Process and Rationale," NASA Engineering and Safety Center Technical Report RP-08-118, Nov. 20, 2008, URL: http://www.nasa.gov/offices/nesc/reports/index.html.

[7]Babula, M., Bertch, W. J., Green, L. L., Hale, J. P., Mosier, G. E., Steele, M. J., and Woods, J., "NASA Standard for Models and Simulations: Credibility Assessment Scale," AIAA Paper 2009-1011, Jan. 2009.

[8]Return to Flight Task Group, "Return to Flight Task Group Final Report," U. S. Government Printing Office, August 2005, URL: http://www.nasa.gov/pdf/125343main_RTFTF_final_081705.pdf, July 2005.

[9]*NASA Governance and Strategic Management Handbook*, NASA NPD 1000.0A, Aug. 1, 2008, URL: http://nodis3.gsfc.nasa.gov/displayDir.cfm?t=NPD&c=1000&s=0A.

[10]"VV&A Recommended Practices Guide, Build 3.0," Defense Modeling and Simulation Office, September 2006, URL: http://vva.dmso.mil/ [cited February 2008].

[11]"Guide for Verification and Validation of Computational Fluid Dynamics Simulations," American Institute of Aeronautics and Astronautics, AIAA G-077, 1998.

[12]"Guide for Verification and Validation in Computational Solid Mechanics," American Society of Mechanical Engineers, ASME V&V 10, New York, NY, 2006.

[13]NASA, "NASA Software Policies," NPD 2820.1B, URL: http://nodis3.gsfc.nasa.gov/.

[14]NASA, "NASA Software Engineering Requirements," NPR 7150.2, URL: http://nodis3.gsfc.nasa.gov/.

[15]Sabra, A. I. (ed.), *The Optics of Ibn al-Haytham, Books I-II-III: On Direct Vision*, The Arabic text, edited and with Introduction, Arabic-Latin Glossaries and Concordance Tables, Kuwait: National Council for Culture, Arts and Letters, 1983.

[16]Grosseteste, R., *Posterior Analytics,* 1235 [see A. C. Crombie, *Robert Grosseteste and the Origins of Experimental Science, 1100-1700*, Oxford: Clarendon Press, 1971, pp. 52-60]

[17]Bacon, R., *Opus maius*, 1267, Edited by J. H. Bridges, *The 'Opus Maius' of Roger Bacon*, Oxford, 1879-90, 3 vols., reprinted 1964, Part III, *De signis (On Signs),* ed. K.M. Fredborg, L. Nielsen and J. Pinborg, 'An Unedited Part of Roger Bacon's "Opus Maius: De Signis"',' *Traditio,* Vol. 34, pp. 75-136, 1978.

[18]NASA, "NASA Systems Engineering Processes and Requirements," NPR 7123.1, URL: http://nodis3.gsfc.nasa.gov/.

[19]IEEE, *IEEE standard dictionary of electrical and electronics terms*, ANSI/IEEE Std 100-1984, 1984.

[20]Oberkampf, W. L. and Trucano, T. G., "Verification and validation in computational fluid dynamics," *Progress in Aerospace Sciences,* Vol. 38, 2002, pp. 209-272.

[21]Balci, O., "A Methodology for Certification of Modeling and Simulations Applications," *ACM Transactions on Modeling and Computer Simulation*, Vol. 11, No. 4, October 2001, pp. 352-377.

[22]Department of Defense, "DoD Modeling and Simulation (M&S) Verification, Validation, and Accreditation (VV&A)," DoD Instruction 5000.61, 2003.

[23]NASA, "Risk Management Procedural Requirements," NPR 8000.4, URL: http://nodis3.gsfc.nasa.gov.