Credibility Theme Strategic Insight 2018

ASSESS

UNDERSTANDING AN ENGINEERING SIMULATION RISK MODEL

Acknowledgements

The ASSESS Initiative would like to acknowledge the Engineering Simulation **Credibility** Working Group for its efforts to support advancement of this ASSESS theme and the development of this Strategic Insight Paper along with the efforts of those who participated in Working Sessions at the ASSESS annual Congresses on the **Credibility** Theme. The ASSESS Initiative would also like to acknowledge the excellent work done in the NASA Standard for Models and Simulations (NASA-STD-7009A) and the Sandia Predictive Capability Maturity Model (PCMM).

The ASSESS Initiative would like to acknowledge specifically the following individuals for their efforts on this ASSESS Strategic Insight Paper.

Primary Author:

Joe Walsh, ASSESS Initiative/intrinSIM

Reviewers:

Rod Dreisbach, Engineering Simulation Consultant
Keith Meintjes, CIMdata
Tina Morrison, US FDA
Martin Steele, NASA
Walter Witkowski, Sandia National Laboratories

Contents

ENGINEERING SIMULATION RISK MODEL

Engineering Simulation is being used more and more broadly to make informed technical and business decisions, especially during the early stages of developing a new product. Use of an Engineering Simulation Risk Model improves credibility through a clearer understanding of the predictive capabilities and "appropriateness" of the simulation(s), thereby increasing confidence in Engineering Simulation influenced decisions.

Engineering Simulation as defined by NAFEMS is "The use of numerical, physical or logical models of systems and scientific problems in predicting their response to different physical conditions."

What is an Engineering Simulation Risk Model?

An Engineering Simulation Risk Model (ESRM) is a Predictive Capability Assessment that includes a set of recommended best practices and associated metrics to understand and manage the "appropriateness" and risk associated with Engineering Simulation influenced decisions.

The approach outlined in this paper for an ESRM is founded on the concept of determining the "appropriateness" of the simulations performed to influence/support the decision(s) being made. This approach to Predictive Capability Assessment is inspired by the NASA-STD-7009A "Use Assessment", Sandia Predictive Capability Maturity Model, the emerging ASME V&V 40 standard, NAFEMS Simulation Governance and Management Working Group activities, and the ASSESS Credibility Theme discussions.

An ESRM is independent of the type of simulation performed and is intended to account for accuracy, sensitivity, and uncertainty of the simulation along with the criticality and risk associated with the decision(s) being made. The ASSESS initiative has repeatedly stated that, "Every Engineering Simulation that is performed should be done to support either a decision to be made or a decision that has been made." Therefore, any ESRM must account for the intended use of the simulation being performed.

This ASSESS Strategic Insight Paper starts with a brief review of key efforts in this area at NASA (NASA-STD-7009A) and Sandia (Predictive Capability Maturity Model). Following this, an approach for a generalized ESRM providing a set of recommendations and criteria that could be developed and used in support of Engineering Simulation influenced decisions. The proposed approach includes guidelines, and a potential framework for defining an ESRM. Each organization will need to assess and then tailor the proposed guidelines and framework when implementing an ESRM for their specific applications.

The objective of this ASSESS Strategic Insight Paper is to initiate discussions leading to further development of an effective generalized Engineering Simulation Risk Model that can be used to support the current and dramatically expanding use of Engineering Simulation. Understanding and enabling an effective and consistent ESRM is one key to enabling a significantly broader use of Engineering Simulation in support of more informed business and technical decision making.

An Engineering Simulation Risk Model is a Predictive Capability Assessment that includes a set of recommended best practices and associated metrics to understand and manage the "appropriateness" and risk associated with Engineering Simulation influenced decisions.

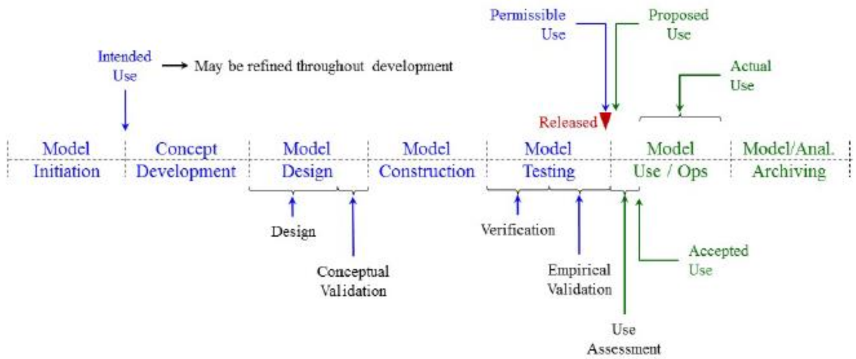
NASA-STD-7009A STANDARD FOR MODELS AND SIMULATIONS

The NASA-STD-7009A document outlines requirements and recommendations for modeling and simulation (M&S) activities within NASA. The primary purpose of NASA-STD-7009A is to reduce the risks associated with M&S-influenced decisions by ensuring a more complete communication of M&S results, including assessments of criticality and credibility. This NASA Technical Standard establishes practices to help reduce, assess, and communicate risk by making the factors leading to M&S credibility more apparent through the use of assessments that occur during different stages of the life cycle of M&S.

NASA-STD-7009A establishes practices to help reduce, assess, and communicate risk by making the factors leading to M&S credibility more apparent.

The M&S Lifecycle at NASA

NASA-STD-7009A defines life cycle of a model or simulation, like that of any system, as being comprised of two general parts: M&S development, which includes M&S initiation, M&S concept development, M&S design, M&S construction, M&S testing, and M&S application which includes M&S use. These phases are illustrated in the following figure from NASA-STD-7009A with different "uses" called out.



The following definitions of "uses" and use assessment are provided in NASA-STD-7009A as shown below.

- Intended Use: The expected purpose and application of the M&S.
- Permissible Use: The purposes for which the M&S is formally allowed.
- Proposed Use: A desired specific application of the M&S.
- Use Assessment: The process of determining if the M&S is accepted for a Proposed Use.
- Accepted Use: The successful outcome of a Use Assessment designating the M&S is accepted for a Proposed Use.
- Actual Use: The specific purpose and domain of application for which the M&S is being, or was, used.

The first step of M&S is to define clearly the purpose of the model, how the model is to be used, and the expectations of the simulation results including accuracy and uncertainty quantifications. The results of this M&S planning step lead to a definition of the intended use of the M&S.

The use of a model starts with an assessment of whether the Proposed Use of the model sufficiently matches the Permissible Use. If the Proposed Use is acceptable, the model is then used to obtain the results of interest. If the Proposed Use does not meet the defined Permissible Use, the Proposed Use should either be rejected or possibly allowed with the appropriate restrictions, caveats, or placarding. Each application of the model restarts the model use/operation with Use Assessment of the needs of a specific Proposed Uses against the Permissible Uses.

The use of a model starts with an assessment of whether or not the Proposed Use of the model sufficiently matches the Permissible Use.

The credibility of the results from the Actual Use of a model is assessed using the Credibility Assessment requirements (described later).

NASA -STD-7009A further recommends that the responsible party shall:

1. Perform and document the Criticality Assessment for the planned M&S activity.
2. Define the objectives and requirements for M&S activities including:
    a. Acceptance criteria
    b. Intended Use
    c. Metrics
    d. Verification, validation, and uncertainty characterization
    e. Reporting of M&S information for critical decisions
    f. Configuration management
3. Document any technical reviews accomplished in regard to the development, management (control), and use of the M&S

Criticality Assessment

Decisions based entirely or partially on M&S are usually made within the context of a program or project. The risk assumed by the program or project is often incorrectly estimated due to several factors that may occur throughout the development and use of an M&S, not the least of which is an inadequate assessment of uncertainties.

A Criticality Assessment provides a proactive method to understand and mitigate potential risks in decision making as early as possible by considering: (a) the consequences to human safety, other risks, product design criteria or business criteria, and (b) the degree to which M&S results are used to influence a decision.

Decision Consequence

Decision Consequence classifications assess the impact of an M&S-influenced decision. NASA-STD-7009A outlines that the Decision Consequence be evaluated across nine criteria:

1. Personnel
2. Flight or Ground Hardware
3. Flight or Ground Equipment
4. Flight or Ground Facilities
5. Operational Status
6. Capabilities
7. Schedules
8. Cost
9. Mission Success Criteria.

NASA-STD-7009A outlines five classifications for each Decision Consequence classification and the criteria for determining the appropriate classification associated with each of the nine criteria listed above.

Decision Consequence classifications:

I. Negligible
II. Minor
III. Moderate
IV. Significant
V. Catastrophic

M&S Influence

M&S Influence estimates the degree to which M&S results might impact the decision under consideration. This is predicated on the amount of other information available when making the impending decision. NASA-STD-7009A outlines three criteria to be considered when evaluating M&S Influence:

1. Real systems in a real environment
2. Similar systems in a similar environment
3. Other Engineering Simulation(s).

It also outlines five classifications for each M&S Influence and the criteria for determining the appropriate classification associated with each of the three criteria listed above.

M&S Influence classifications:

1. Negligible
2. Minor
3. Moderate
4. Significant
5. Controlling

A Criticality Assessment considers (a) the consequences to human safety, other risks, product design criteria or business criteria, and (b) the degree to which M&S results are used to influence a decision.

Criticality Assessment Matrix

A Criticality Assessment Matrix may be developed by combining the Decision Consequence on one axis with the Engineering Simulation Influence on the other axis. This Matrix is used by NASA-STD-7009A to determine the need for a Credibility Assessment.

The following figure illustrates the Engineering Simulation Criticality Assessment Matrix with a stoplight (Red, Yellow, Green) classification of Criticality.

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

From the perspective of situational criticality, the three possible cases for assessment are:

1. Those M&S that are assessed to fall within the red (R) boxes in are clear candidates for fully following NASA-STD-7009A.
2. Those M&S that are assessed to fall within the yellow (Y) boxes may or may not be candidates for fully following NASA-STD-7009A at the discretion of program/project management in collaboration with the Technical Authority.
3. Those M&S that fall within the green (G) boxes have full discretion for following NASA-STD-7009A.

A Criticality Assessment Matrix may be developed by combining the Decision Consequence on one axis with the Engineering Simulation Influence on the other axis.

Credibility Assessment Factors

The credibility of M&S-based results is not something that can be assessed directly; however, key factors of credibility may be assessed more directly.

Eight factors were selected from a long list of factors that potentially contribute to the credibility of M&S results because (a) they were individually judged to be the key factors in this list; (b) they are nearly orthogonal, i.e., largely independent from one another; and (c) they can be assessed objectively. These eight Credibility Assessment Factors are grouped into the three categories of Development (Data Pedigree, Verification, Validation); Operations (Input Pedigree, Uncertainty Characterization, Results Robustness); and Supporting Evidence (M&S History, M&S Process/Product Management.

Development			Operations			Supporting Evidence	
Data Pedigree	Verification	Validation	Input Pedigree	Uncertainty Characterization	Results Robustness	M&S History	M&S Process/Product Mgt

The M&S Development category captures those aspects of the M&S that pertain to the purposes for which it was developed; the M&S Use (Operations) category addresses the aspects relevant to the current application of the M&S to generate the particular M&S results under assessment; and the Supporting Evidence category addresses two cross-cutting factors related to history & process management of the M&S activity.

The NASA-STD-7009A further defines levels from 0 to 4 for ranking each of the Credibility Assessment Factors. For any level, when multiple conditions are stated, all of them should be met to achieve that level unless they are expressed as an "or" condition. The assessment of each factor is a discrete step function, with no intentions for partial credit at any given level. The NASA-STD-7009A clearly states that the factor assessments are ordinal only, and as such, should not be arithmetically manipulated.

Data Pedigree Factor

Data Pedigree involves the evaluation of all data used in the development of the M&S, and is formally defined as a record of the traceability of data from its source through all aspects of its transmission, storage, and processing to its final form used in the development of the M&S. Any changes from the source data may be of significance to its pedigree. Ideally, this record includes important quality characteristics of the data at every stage of the process.

The credibility of M&S-based results is not something that can be assessed directly, however, key factors of credibility may be assessed more directly. Eight factors were selected from a long list of factors that potentially contribute to the credibility of M&S results.

NASA-STD-7009A specifies the following criteria for ranking for the Data Pedigree Factor.

Data Pedigree Factor Criteria	
Level	**Status**
0	Insufficient evidence.
1	Some data is known and traceable to informal documentation. Sources of all significant data are known. Uncertainties in data may not even be estimated.
2	Most data are known and are traceable to formal documentation. Processes to establish significant data are known. Uncertainties in all data are at least estimated.
3	All data are known and can be traced to a sufficiently representative referent. All significant data are acceptable in terms of accuracy, precision and uncertainty.
4	All data known and fully traceable to the real-world system. All data are acceptable in terms of accuracy, precision, and uncertainty.

Verification Factor

Verification is the process used to provide assurance that the computational model is implemented correctly. There are two different aspects with respect of verification: (a) code verification and (b) solution verification.

Code verification employs standard software development techniques, including regression testing and unit testing. Ultimately, code verification should be accomplished via the use of the end-to-end computational model to ensure interactions between the model components are correct.

Solution verification involves identifying the presence of any numerical and logical errors in the model, assessing their impact upon the accuracy of the results, and taking necessary steps (if any) to ensure minimal adverse impacts on any of the simulation requirements.

Other key aspects to consider include: (a) the degrees of rigor and formality of the verification processes, and (b) how well-established and appropriate the processes are in the context of the specific M&S being developed.

NASA-STD-7009A specifies the following criteria for the Verification Factor.

Verification is the process used to provide assurance that the computational model is implemented correctly. There are two different aspects with respect of verification: (a) code verification and (b) solution verification.

Verification Factor Criteria	
Level	**Status**
0	Insufficient evidence.
1	Verification is informal, with some documentation or evidence of completeness/success.
2	The model is correctly implemented as determined by documented verification practices, which evaluate all components, features, capabilities, and couplings of the model. Documented methods are used to assess model errors. Most of the important model errors satisfy program/project-specified requirements.
3	The model is correctly implemented as determined by formal verification practices, which evaluate all components, features, capabilities, and couplings of the end-to-end model. Formal methods are used to assess model errors. All important model errors satisfy program/project-specified requirements.
4	The model is correctly implemented as determined by reliable verification practices, which evaluate all components, features, capabilities, and couplings of the end-to-end model. Reliable estimation methods are used to assess model errors. All model errors satisfy program/project specified requirements.

Validation starts with providing the requisite assurance that the conceptual and mathematical models are valid. Once the computation model is available, the next step is empirical validation, which is the comparison of M&S results with a referent.

Validation Factor

Validation starts with providing the requisite assurance that the conceptual and mathematical models are valid (i.e., Conceptual Validation). Once the computational model is available, the next step is empirical validation, which is the comparison of M&S results with a referent. In some instances, the referent can be the results obtained from a higher-fidelity model.

The Validation factor considers the following aspects when assessing credibility: (a) the similarity between the real-world system and the referent, (b) the extent of the domain of M&S validation relative to the domain of the real-world system operation, and (c) the extent to which favorable comparison is achieved for all possible model outputs. Specific criteria must be defined for what constitutes “favorable comparison.” Each model output may have a unique domain of validation, i.e., favorable comparison may not be obtained for all model outputs for each set of model inputs.

The comparison between M&S results and the referent data must consider: (a) the accuracy of the results – for computational models, the magnitude of the numerical difference between the mean of the M&S result and the mean of the referent data, and (b) the associated uncertainty, i.e., the spread about the means. The comparison may also include sensitivities of the results with respect to corresponding independent variables in both model and experiment.

NASA-STD-7009A specifies the following criteria for ranking for the Validation Factor.

Validation Factor Criteria	
Level	**Status**
0	Insufficient evidence.
1	The model is conceptually validated. The problem statement (intended use) is clearly stated & well-understood, and the conceptual model, requirements, & specifications are correct and sufficiently address the problem.
2	M&S results compare favorably to measurements from a representative system or to results from a higher-fidelity M&S that satisfies the conditions for Level 2. Validation points are within the domain of operation for the Real-World System (RWS). Favorable comparisons are obtained for at least some of the important response quantities.
3	M&S results compare favorably to measurements on the RWS in a representative environment or to results from a higher-fidelity M&S that satisfies the conditions for Level 3. Validation points significantly span the domain of operation for the RWS. Favorable comparisons are obtained for all important response quantities.
4	M&S results compare favorably to measurements on the RWS in its operating environment or to results from a higher-fidelity M&S that satisfies the conditions for Level 4. Validation points completely span the domain of operation for the RWS. Favorable comparisons are obtained for all response quantities.

Input Pedigree Factor

Input Pedigree involves the evaluation of all data used as input to the M&S. It is formally defined as a record of the traceability of data from its source through all aspects of its transmission, storage, and processing to its final form when using the M&S. Any changes from the source data may be of significance to its pedigree. Ideally, this record includes important quality characteristics of the data at every stage of the process.

NASA-STD-7009A specifies the following criteria for ranking for the Input Pedigree Factor.

Input Pedigree Factor Criteria	
Level	**Status**
0	Insufficient evidence.
1	Some input data are known and traceable to informal documentation. Sources of all significant input data are known. Uncertainties in input data may not even be estimated.
2	Most input data are known and traceable to formal documentation. Processes to establish significant data are known. Uncertainties in all data are at least estimated.
3	All input data are known and can be traced to a sufficiently representative referent. All significant data are acceptable in terms of accuracy, precision, and uncertainty.
4	All input data known and fully traceable to the RWS. All data are acceptable in terms of accuracy, precision, and uncertainty.

Input Pedigree involves the evaluation of all data used as input to the M&S. It is formally defined as a record of the traceability of data from its source through all aspects of its transmission, storage, and processing to its final form when using the M&S.

Uncertainty Characterization Factor

Uncertainty Characterization includes the identification of uncertainty sources and the qualification or quantification of uncertainty in the current M&S results. Important aspects of Uncertainty Characterization are (a) the sources of uncertainty in the input variables and parameters, (b) the numerical errors incumbent in model implementation mechanisms (e.g., computational/math models), and (c) the propagation of the uncertainty to M&S outputs.

NASA-STD-7009A specifies the following criteria for ranking for the Uncertainty Characterization Factor.

Uncertainty Characterization Factor Criteria

Level	Status
0	Insufficient evidence.
1	Sources of input uncertainty have been identified with qualitative estimates of the uncertainty. Their impact on output uncertainties and uncertainty propagation has not been addressed.
2	Most sources of uncertainty have been identified, expressed quantitatively and correctly classified based on Subject Matter Expert (SME) opinions and/or by deduction from experimental data. Propagation of the uncertainty to output quantities has been addressed by reduced order and/or reduced dimension propagation.
3	Quantitative estimates of uncertainties have been reported for most output quantities after a propagation of all known sources of uncertainty.
4	A statistical analysis of the output uncertainty has been performed for all output quantities after rigorous and validated propagation of all known sources of uncertainty. Reported results may include statistical moments, confidence intervals, sensitivity analysis, etc..

Results Robustness Factor

Results Robustness is the determination of how thoroughly the sensitivities of the current M&S results are known. Simulations aim to imitate the real world or a proposed real world. Ideally, the imitated system behaves like the RWS (i.e., with acceptable accuracy and precision, and with similar sensitivities). If the RWS is sensitive to certain variables or parameters, then the M&S results should be similarly sensitive.

Sensitivity Analysis determines the stability (robustness) of the scenario(s) under analysis. The key sensitivities are defined as parameters and variables shown to produce large changes in results with relatively small perturbations to input.

NASA-STD-7009A specifies the following criteria for ranking for the Results Robustness Evidence Factor.

Uncertainty Characterization includes the identification of uncertainty sources and the qualification or quantification of uncertainty in the current M&S results.

Results Robustness Factor Criteria	
Level	**Status**
0	Insufficient evidence.
1	Sensitivity of M&S results for the RWS is estimated by analogy with the quantified sensitivity of similar problems of interest.
2	Sensitivity of the M&S results for the RWS is quantitatively known for a few variables and parameters. Only a few of the most sensitive variables and parameters are identified. Sensitivities of combinations of variables and parameters are not known.
3	Sensitivity of the M&S results for the RWS is quantitatively known for many variables and parameters, including many of the most sensitive variables and parameters. Sensitivities of some combinations of these variables and parameters are also quantified.
4	Sensitivity of the M&S results for the RWS is quantitatively known for most of the variables and parameters, including most, if not all, of the most sensitive variables and parameters. Sensitivities of many combinations of these variables and parameters are also quantified.

Results Robustness is the determination of how thoroughly the sensitivities of the current M&S results are known.

M&S History Factor

The M&S History implicitly includes two main elements: change history and use history. This factor provides an assessment of the "heritage" of the M&S from these two viewpoints.

1. The change history sub-factor assesses the degree of changes of the current model relative to versions used in previous applications.
2. The use history sub-factor assesses the degree to which the current proposed or actual use of the M&S is identical to previous uses.

NASA-STD-7009A specifies the following criteria for ranking for the M&S History Factor.

M&S History Factor Criteria	
Level	**Status**
0	Insufficient evidence.
1	Model is new or has major changes from previously used versions, or proposed use has major differences from previous uses; however, the model, changes, and uses are documented.
2	Model has at most moderate changes from previously used versions, and proposed use is at most (no more than) moderately different from previous uses.
3	Model has at most minor changes from previously used versions, and proposed use is in interpolated regions of validation points.
4	Model changes have stabilized to inconsequential levels and proposed use is within established norms for the model.

M&S Process/Product Management Factor

The term M&S Process/Product Management is used to describe the extent to which an M&S effort exhibits the characteristics of work product management; process definition; process measurement; process control; process change; continuous improvement, including Configuration Management (CM); and M&S support and maintenance. This factor assesses how rigorously the processes and products of an M&S are managed and maintained.

NASA-STD-7009A specifies the following criteria for ranking for the M&S History Factor.

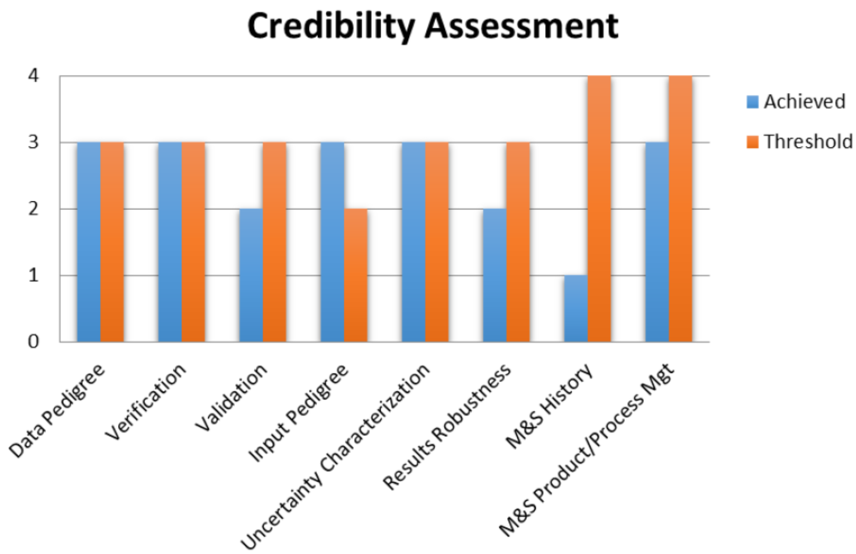
M&S Process/Product Management Factor Criteria	
Level	**Status**
0	Insufficient evidence.
1	Roles and responsibilities are defined in the context of an M&S process that is informally documented. Requirements for M&S products are informally documented. CM of M&S products is established and applied using informal methods.
2	Roles and responsibilities are defined in the context of an M&S process that is formally documented and approved. Requirements for M&S products are formally documented and approved. CM of M&S products is established and applied using formal methods.
3	The formally established process is rigorously controlled and followed. Compliance with the process is formally documented. Measurements of process and product compliance are made and documented. CM of M&S products is rigorously applied.
4	Measurements, including customer/user feedback, are used to improve both the M&S process and products.

Credibility Assessment Sufficiency Thresholds

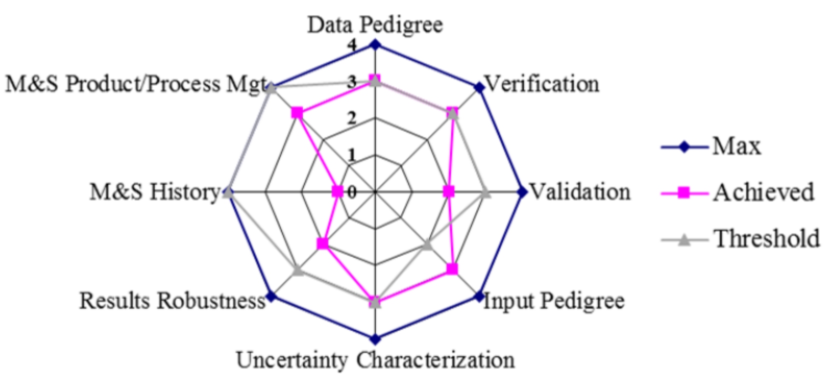
Program/project management, technical authorities, stakeholders, and customers are encouraged to set credibility thresholds (goals) that an M&S effort is to meet. During M&S planning, development, and use, the developers and users can then allocate the appropriate amount of effort to achieving those thresholds. Reporting the factor assessments with their associated thresholds provides an additional basis for credibility discussions.

The term M&S Process/Product Management is used to describe the extent to which the M&S effort exhibits the characteristics of work product management; process definition; process measurement; process control; process change; continuous improvement, including Configuration Management (CM); and M&S support and maintenance.

Reporting the factor assessments with their associated thresholds provides an additional basis for credibility discussions.



Spider Plot (Radar Chart) with Thresholds



Note, as illustrated in the example plots, it is possible that the threshold might not be met, or it might be exceeded.

Predictive Capability Maturity Model (PCMM)

The Predictive Capability Maturity Model (PCMM) developed by Sandia National Laboratories is designed to characterize, plan, and communicate completeness of the approaches used for computational model definition, verification, validation, and uncertainty quantification associated with an intended application. The PCCM is coupled with a Phenomena Identification and Ranking Table (PIRT), peer review, and other pre-requisites as gatekeepers for M&S Activities.

PIRT is an approach developed by the U.S. NRC.

The PCMM was initially developed in 2007 to provide more structure and formality in assessing the maturity and credibility of a computational simulation analysis for a target application, to reduce the ambiguity in such assessments, and to provide specificity as to what should be assessed and communicated to the analyst's customer. The PCCM at Sandia National Laboratories undergoes continuous review and updating.

The PCMM was proposed as a structured method for:

a. planning and organizing M&S Efforts
b. assessing the maturity level of M&S efforts
c. communicating the maturity level of M&S efforts.

Maturity Model Levels

The PCMM is primarily interested in providing M&S maturity assessment information for an application of interest to program managers, interested stakeholders, and decision makers. Desirable goals for a PCMM implementation include:

(1) consistent evaluation
(2) consistent communication
(3) consistent application of the evaluation results.

The PCMM maturity scale assesses the maturity of the M&S effort, or process, directed toward an engineering system of interest but does not by itself assess whether the M&S effort, the accuracy of the predictions, or the performance of the engineering system satisfies a set of imposed requirements. The PCCM is coupled with a PIRT, peer review, and other pre-requisites as gatekeepers for M&S Activities.

The PCMM outlines the general characteristics of four levels of maturity ranging from 0 to 3 that apply to the following M&S elements:

Maturity Level Characteristics

Level	Status
0	Little or no assessment of the accuracy or completeness has been made; little or no evidence of maturity; individual judgment and experience only; convenience and expediency are the primary motivators.
1	Some informal assessment of the accuracy and completeness has been made; generalized characterization; some evidence of maturity; some assessment has been made by an internal peer review group.
2	Some formal assessment of the accuracy and completeness has been made; detailed characterization; significant evidence of maturity; some assessments have been made by an internal peer review group.
3	Formal assessment of the accuracy and completeness has been made; precise and accurate characterization; detailed and complete evidence of maturity; essentially all assessments have been made by independent peer-review groups.

The Predictive Capability Maturity Model (PCMM) developed by Sandia National Laboratories is designed to characterize, plan, and communicate completeness of the approaches used for computational model definition, verification, validation, and uncertainty quantification associated with an intended application.

An overall understanding of what the customer requires is necessary to identify what physics need to be modeled, and the desired level of fidelity in modeling the physics leading to the quantities of interest. Because the evaluation of the PCMM is for an intended application (or family of intended applications), knowledge of what this application entails, and what the customer expects from a computational simulation, are necessary. A PIRT for the customer application identifies and rates the importance of the relevant physical phenomena that are required to model the customer's application and performs a high-level gap analysis on how well this phenomena/physics can be modeled using a selected computational simulation tool.

M&S Elements

Six M&S elements were identified as contributors to the M&S process and used to assess maturity in the PCMM model are:

(1) Representation and Geometric Fidelity
(2) Physics and Material Model Fidelity
(3) Code Verification
(4) Solution Verification
(5) Model Validation
(6) Uncertainty Quantification and Sensitivity Analysis.

Representation and Geometric Fidelity

Representation and geometric modeling fidelity refer to the level of detail included in the spatial definition of all constituent elements of the system being modeled and analyzed. A system means any engineered or natural system entity, e.g., a subsystem, a component, or a part of a component. In M&S, the representational and geometrical definition of a system are commonly specified in a computer-aided design or computer-aided manufacturing (CAD/CAM) software package.

A key issue that complicates the mapping of CAD/CAM geometries to a geometry that is ready for constructing a computational model is that the mapping is dependent on the particular type of physics to be modeled and the specific assumptions in the modeling. For example, a change in material properties along the surface of a missile would be important to a structural dynamics analysis, but it may not be important to an aerodynamic analysis.

The PCMM specifies four Maturity Levels for the Representation and Geometric Fidelity Element, based on the criteria noted below.

A PIRT for the customer application identifies and rates the importance of the relevant physical phenomena that is required to model the customer's application and performs a high-level gap analysis on how well this phenomena/physics can be modeled using a selected computational simulation tool.

Representation and Geometric Fidelity	
Level	**Status**
0	• Judgment only • Little or no representational or geometric fidelity for the system and Boundary Conditions
1	• Significant simplification or stylization of the system and Boundary Conditions • Geometry or representation of major components is defined
2	• Limited simplification or stylization of major components and Boundary Conditions • Geometry or representation is well defined for major components and some minor components • Some peer review conducted
3	• Essentially no simplification or stylization of components in the system and Boundary Conditions • Geometry or representation of all components is at the detail of "as built", e.g., gaps, material interfaces, fasteners • Independent peer review conducted

Representation and geometric modeling fidelity refer to the level of detail included in the spatial definition of all constituent elements of the system being analyzed.

Physics and Material Model Fidelity

The range of physics modeling fidelity can range from empirical models that are based on the fitting of experimental data (empirical models) to what is typically called "first-principles physics." An important aspect of physics modeling fidelity is the degree to which various types of physics are included and coupled in the mathematical model of the system and the environment.

The PCMM specifies four Maturity Levels for the Representation and Geometric Fidelity Element, based on the criteria noted below.

An important aspect of physics modeling fidelity is the degree to which various types of physics are included and coupled in the mathematical model of the system and the environment.

Physics and Material Model Fidelity	
Level	**Status**
0	• Judgment only • Model forms are either unknown or fully empirical • Few, if any, physics informed models • No coupling of models
1	• Some models are physics based and are calibrated using data from related systems • Minimal or ad hoc coupling of models
2	• Physics-based models for all important processes • Significant calibration needed using separate effects tests (SETs) and integral effects tests (IETs) • One-way coupling of models • Some peer review conducted
3	• All models are physics based • Minimal need for calibration using SETs and IETs • Sound physical basis for extrapolation and coupling of models • Full, two-way coupling of models • Independent peer review conducted

Code Verification

The major goal of numerical algorithm verification is to accumulate evidence that demonstrates that the numerical algorithms in the code are implemented correctly and functioning as intended, i.e., they produce the expected convergence rate and correct solution to the specific mathematics being tested.

The PCMM suggests that it is useful to segregate code verification into two activities, numerical algorithm verification and code verification through the use of Software Quality Engineering (SQE) methodology. Numerical algorithm verification addresses the mathematical correctness in the software implementation of all the numerical algorithms that affect the numerical accuracy of the computational results. The emphasis in SQE is on determining whether the code, as part of a software system, is reliable (implemented correctly) and produces repeatable results on specified computer hardware and in a specified computing environment.

The PCMM specifies the following criteria for the four Maturity Levels for the Code Verification Element.

Code Verification	
Level	**Status**
0	• Judgment only • Minimal testing of any software elements • Little or no Software Quality Engineering (SQE) procedures specified or followed
1	• Code is managed by SQE procedures • Unit and regression testing conducted • Some comparisons made with benchmarks
2	• Some algorithms are tested to determine the observed order of numerical convergence • Some features & capabilities (F&C) are tested with benchmark solutions • Some peer review conducted
3	• All important algorithms are tested to determine the observed order of numerical convergence • All important F&Cs are tested with rigorous benchmark solutions • Independent peer review conducted

Solution Verification

Solution verification commonly focuses on the quantitative estimation of the numerical accuracy relative to a given solution to a physics equation. The primary numerical errors that are estimated in solution verification are due to (1) spatial and temporal discretization of PDEs and (2) iterative solution error resulting from a linearized solution approach to a set of nonlinear, coupled equations.

The PCMM specifies four Maturity Levels for the Solution Verification Element, based on the criteria noted below.

The PCMM Model suggests that it is useful to segregate code verification into two activities, numerical algorithm verification and code verification.

Solution verification commonly focuses on the quantitative estimation of the numerical accuracy of a given solution to a physics equation chosen in M&S.

Solution Verification	
Level	**Status**
0	• Judgment only • Numerical errors have an unknown or large effect on simulation results
1	• Numerical effects on relevant System Response Quantities (SRQs) are qualitatively estimated • Input/output (I/O) verified only by the analysts
2	• Numerical effects are quantitatively estimated to be small on some SRQs • I/O independently verified • Some peer review conducted
3	• Numerical effects are determined to be small on all important SRQs • Important simulations are independently reproduced • Independent peer review conducted

Model Validation

Model Validation refers to assessing the accuracy of the computational model and the input data by comparing the System Response Quantities (SRQs) of interest with experimentally measured SRQs. This Element Focuses on:

- Thoroughness and precision of the accuracy assessment of the computational results relative to the experimental measurements
- Completeness and precision of the characterization of the experimental conditions and measurements
- Relevancy of the experimental conditions, physical hardware, and measurements in the validation experiments compared to the application of interest

The PCMM specifies four Maturity Levels for the Model Validation Element, based on the criteria noted below.

Model Validation	
Level	**Status**
0	• Judgment only • Few, if any, comparisons with measurements from similar systems or applications
1	• Quantitative assessment of accuracy of SRQs not directly relevant to the application of interest • Large or unknown experimental uncertainties
2	• Quantitative assessment of predictive accuracy for some key SRQs from IETs and SETs • Experimental uncertainties are well characterized for most SETs, but poorly known for IETs • Some peer review conducted
3	• Quantitative assessment of predictive accuracy for all important SRQs from IETs and SETs at conditions/geometries directly relevant to the application • Experimental uncertainties are well characterized for all IETs and SETs • Independent peer review conducted

Model Validation refers to assessing the accuracy of the computational model and the input data by comparing the System Response Quantities (SRQs) of interest with experimentally measured SRQs.

Uncertainty Quantification and Sensitivity Analysis

Uncertainty Quantification and Sensitivity Analysis refers to the characterization of model predictive uncertainty primarily deals with the proper estimation and representation of all uncertainties that could exist as part of the prediction for the system of interest. This Element focuses on:

- Thoroughness and soundness of the uncertainty quantification effort, including identification and characterization of all plausible sources of uncertainty
- Accuracy and correctness of propagating uncertainties through a computational model and interpreting uncertainties in the SRQs of interest
- Thoroughness and precision of a sensitivity analysis to determine the most important contributors to uncertainty in system responses

Recognition of uncertainties refers to the activity of identifying and understanding all possible uncertainties within the system of interest, such as parametric uncertainty and uncertainties in the geometry, BCs, forcing functions, or environmental conditions. The important issue with Uncertainty Quantification is not the SRQs per se but the estimated total uncertainty in the SRQs of interest as a function of the input parameters and conditions that could exist over the domain of the intended use.

The PCMM specifies four Maturity Levels for the Uncertainty Quantification and Sensitivity Analysis Element based on the criteria noted below.

Uncertainty Quantification and Sensitivity Analysis	
Level	**Status**
0	• Judgment only • Only deterministic analyses are conducted • Uncertainties and sensitivities are not addressed
1	• Aleatory and epistemic (A&E) uncertainties propagated, but without distinction • Informal sensitivity studies conducted • Many strong UQ/SA assumptions made
2	• A&E uncertainties segregated, propagated and identified in SRQs • Quantitative sensitivity analyses conducted for most parameters • Numerical propagation errors are estimated, and effects known • Some strong assumptions made • Some peer review conducted
3	• A&E uncertainties comprehensively treated and properly interpreted • Comprehensive sensitivity analyses conducted for parameters and models • Numerical propagation errors are demonstrated to be small • No significant UQ/SA assumptions made • Independent peer review conducted

Uncertainty Quantification and Sensitivity Analysis refers to the characterization of model predictive uncertainty primarily deals with the proper estimation and representation of all uncertainties that could exist as part of the prediction for the system of interest.

Maturity Model Requirements

The PCMM states the essential question to ask for each Element of the maturity model is, "what should the appropriate level of maturity be for my intended use of the M&S activity." It also includes a general reference for applicability / requirements of maturity levels in its definition of the maturity level characteristics as outlined in the table below.

Maturity Level	Consequence	M&S Impact	Example
0	Low	Minimal	Scoping Studies
1	Moderate	Some	Design Support
2	High	High	Qualification Support
3	High	Basis for Decision	Qualification or Certification support

The following table depicts example results of specifying project maturity requirements for each of the assessed elements. The designator "Required" is used to indicate the project maturity requirement for each element.

The values in the table are color coded to designate the following:

- Green – The assessment meets or exceeds the requirement.
- Yellow – The assessment does not meet the requirement by one level or less.
- Pink – The assessment does not meet the requirement by two levels or less.
- Red – The assessment does not meet the requirement by three levels or less.

MATURITY / ELEMENT	Maturity Level 0	Maturity Level 1	Maturity Level 2	Maturity Level 3
Representation and Geometric Fidelity		Assessed	Required	
Physics and Material Model Fidelity			Assessed Required	
Code Verification		Assessed Required		
Solution Verification	Assessed		Required	
Model Validation		Assessed	Required	
Uncertainty Quantification and Sensitivity Analysis	Assessed			Required

The PCMM maturity scale assesses the maturity of an M&S effort, or process, directed toward an engineering system of interest. The scale, by itself, does not assess whether the M&S effort, the accuracy of the predictions, or the performance of the engineering system satisfies a set of imposed requirements.

It is recommended that the M&S related team has the flexibility in defining the desired levels required for each Element and that there is no requirement for the desired levels to be the same for all elements. However, PCMM does require that the maturity level of the elements Model Validation and Uncertainty Quantification and Sensitivity Analysis should be no higher than two levels above the maturity levels of the minimum of code verification and solution verification. This requirement means that the maturity levels of code verification and solution verification must be assessed before the maturity levels of model validation and of uncertainty quantification and sensitivity analysis are assessed.

The PCMM further mentions that M&S reviews typically require some type of compression of the PCMM scoring. A summary method that always maintains a minimum value, an average value, and a maximum value through any aggregation process is recommended.

A Generalized Engineering Simulation Risk Model (ESRM)

The primary purpose of the generalized Engineering Simulation Risk Model (ESRM) proposed herein is to improve understanding of the "credibility" of each phase of an Engineering Simulation, thereby increasing confidence in the Engineering Simulation influenced decisions. The secondary purpose is to initiate discussions leading to a potential path to a scalable ESRM that can be deployed broadly across multiple organizations. The approaches outlined in this paper are conceptual in nature, where the full definition of a deployable ESRM is beyond the scope of this paper.

The previously explored efforts of the NASA-STD-7009A and the PCMM model provide excellent insights and associated metrics to understand and manage the risk associated with Engineering Simulation influenced decisions. However, these extraordinary efforts are also designed primarily the Engineering Simulation activities, nomenclature, and processes of NASA and Sandia National Laboratories making their adoption by the Commercial sector challenging.

A generalized ESRM is an attempt to enable effective broader use of Engineering Simulation across multiple organizations, industries, and applications involved in all three Engineering Simulation sectors (Commercial, Academic, Government). The ESRM proposed in this paper is an example of a generalized Predictive Capability Assessment that leverages information and principles from the NASA-STD-7009A and the PCMM model.

A generalized Engineering Simulation Risk Model is an attempt to enable effective broader use of Engineering Simulation across multiple organizations, industries, and applications involved in all three Engineering Simulation sectors (Commercial, Academic, Government).

Engineering Simulation involves multiple activities including; development of underlying algorithms, coding of algorithms into software, methodology and/or process development, methodology and process implementation, instance-specific model definition, simulation execution, and results evaluation. These Engineering Simulation activities can be grouped into three major phases:

1. Algorithm & Software Development
2. Methodology & Process Development
3. Methodology & Process Application

The Credibility Reviews in each phase should align with the activity and is typically done separately with an understanding of the relationship with the other phases. The needs for each phase are:

- Algorithm & Software Development phase: understand potential and permissible uses of the software.
- Methodology & Process Development Phase: understand the “appropriateness” of Algorithms & Code selected and potential and permissible uses of the process.
- Methodology & Process Application Phase: understand the “appropriateness” of the methodology/process selected for use.

This separation into three distinct phases enables a generalized approach where the Criteria Assessment can be performed independently for each phase. Broad applications of Engineering Simulation outside of the Government and Academic sectors are performed primarily using commercial software. It is very rare that the same people are involved in all three phases of activity. The most common Engineering Simulation activities are focused on each phase separately. Previously, simulation experts developed methods and processes and then applied them. The most common (and growing) practice today is that methods and processes are developed by simulation experts which are then used by others.

The proposed concept for a generalized ESRM that is outlined herein is an example of Predictive Capability Assessment that consists of:

a) Usage Impact
b) Phase-Based Predictive Capability Assessment
    a. Determination of Applicable Credibility Reviews
    b. Credibility Objectives
    c. Credibility Reviews
    d. Appropriateness Assessment

This separation into three distinct phases enables a generalized approach where the Criteria Assessment can be performed independently for each phase.

It is very rare that the same people are involved in all three phases of activity.

The proposed Credibility Reviews for each Engineering Simulation phase include:

- Previous Phase (if applicable)
- Usage
- Pedigree (as appropriate to that phase)
- Verification
- Fidelity
- Validation
- Uncertainty
- Robustness

The proposed generalized ESRM is itself a model. This paper is focused on defining and using the Phase-Based Predictive Capability Assessment, where any associated rankings and assessments have not yet gone through any of the foregoing Credibility Reviews.

Usage Impact

Usage Impact as outlined in this paper is a first attempt at generalizing the concepts and approaches outlined in the NASA-STD-7009A Criticality Assessment. This criterion is the same for each of the three phases of Engineering Simulation, however, the application and understanding of this assessment are phase dependent.

Usage Impact considers:

(a) Decision Consequence: the consequences to human safety, other risks, product design criteria or business criteria, and
(b) Engineering Simulation Influence: the degree to which the simulation results might influence a decision

This paper explores Decision Consequence and Engineering Simulation Influence, as well as presenting recommendations on combining them to assess and understand the usage impact of the intended simulation(s).

Decision Consequence

Decision Consequence classifications are used to assess the impact of an Engineering Simulation influenced decision. It is proposed that the following levels of classification, as presented in the NASA-STD-7009A, be used in the ESRM.

1. Negligible
2. Minor
3. Moderate
4. Significant
5. Catastrophic

Usage Impact as outlined in this paper is a first attempt at generalizing the concepts and approaches outlined in the NASA-STD-7009A Criticality Assessment.

The following criteria are proposed for evaluating each Decision Consequence as derived from the Criticality Assessment criteria presented in the NASA-STD-7009A.

- Personal Injury/Harm
- Operational Status
- Capabilities
- Schedules
- Cost
- Business/Mission goals.

It is recommended that the Decision Consequence ranking for any decision should be the highest level of classification across all Decision Consequence criteria as outlined in the following table.

Decision Consequence Ranking	
Criterion	**1. Negligible**
Personal Injury/Harm	Inconsequential.
Operational Status	No effect.
Capabilities	No effect; no degradation.
Schedules	No effect.
Cost	No effect.
Business/Mission Goals	All met.
Criterion	**2. Minor**
Personal Injury/harm	Minor detriment (first aid).
Operational Status	At most a temporary effect.
Capabilities	At most a temporary effect; no more than inconsequential degradation.
Schedules	Minor impact to schedule with no effect on major operational milestones.
Cost	Minor cost impact but within nominal margins.
Business/Mission Goals	All met
Criterion	**3. Moderate**
Personal Injury/harm	Minor injury or occupational illness.
Operational Status	Temporarily off-line for repair.
Capabilities	Temporarily unavailable until restored; some minor degradation.
Schedules	Internal schedule slips with no effect on major milestones.
Cost	Cost overruns beyond nominal margins, but not detrimental to project execution or completion.
Business/Mission Goals	A few not met.

Decision Consequence classifications are used to assess the impact of an Engineering Simulation influenced decision.

Decision Consequence Ranking (continued)	
Criterion	**4. Significant**
Personal Injury/harm	Severe injury or occupational illness.
Operational Status	Permanently degraded until repaired.
Capabilities	Significant or permanent degradation until repaired.
Schedules	Impacts to major mission (operations) milestones.
Cost	Cost overruns detrimental to program or project execution or full completion.
Business/Mission Goals	Many not met.
Criterion	**5. Catastrophic**
Personal Injury/harm	Permanent disability or death.
Operational Status	Non-operational.
Capabilities	Severely degraded to none.
Schedules	Operational (e.g., mission) windows missed.
Cost	Cost overruns cause major program or project reductions or cancellation.
Business/Mission Goals	Most to all not met.

Engineering Simulation Influence refers to the degree to which the results from an Engineering Simulation impact the technical or business decision under consideration.

Engineering Simulation Influence

Engineering Simulation Influence refers to the degree to which the results from an Engineering Simulation impact the technical or business decision under consideration. The definition of an Engineering Simulation ranking is predicated on the amount of other information available when making the impending decision, including determination of whether the design requirements have been verified.

The Engineering Simulation ranking outlined in this paper uses the following levels of classification; the same as presented in NASA-STD-7009A.

1. Negligible
2. Minor
3. Moderate
4. Significant
5. Controlling

The following criteria are proposed for evaluating Engineering Simulation Influence, which are derived from the Criticality Assessment criteria presented in NASA-STD-7009A.

- Real systems in a real environment
- Similar systems in a similar environment
- Other Engineering Simulation(s).

The Engineering Simulation Influence accounts for combinations of states across all three criteria, where the recommended ranking of classifications is outlined in the following table.

Engineering Simulation Influence Ranking				
1. Negligible				
Real System in Real Environment		**Similar System in Similar Environment**		**Other Engineering Simulation**
Ample test data for the real system in the real environment is available.	&	Test data for similar systems in similar environments may or may not be available.	&	No other Engineering Simulation data is used for V&V.
OR				
Some test data for the real system in the real environment is available.	&	Test data for similar systems in similar environments may or may not be available.	&	Credible results from another M&S are also used for V&V.
Ample test data for the real system in the real environment is available.	&	Test data for similar systems in similar environments may or may not be available.	&	No other Engineering Simulation data is used for V&V.
OR				
Some test data for the real system in the real environment is available.	&	Test data for similar systems in similar environments may or may not be available.	&	Credible results from another M&S are also used for V&V.
2. Minor				
Real System in Real Environment		**Similar System in Similar Environment**		**Other Engineering Simulation**
Some test data for the real system in the real environment is available.	&	Test data for similar systems in similar environments may or may not be available.	&	No other Engineering Simulation data is used for V&V.

The definition of an Engineering Simulation ranking is predicated on the amount of other information available when making the impending decision, including determination of whether the design requirements have been verified.

Engineering Simulation Influence Ranking (continued)				
3. Moderate				
Limited test data for the real system in the real environment is available.	&	Ample test data for similar systems in similar environments is available.	&	No other Engineering Simulation data is used for V&V.
OR				
No data is available for the real system in the real environment.	&	Ample test data for similar systems in similar environments is available.	&	Credible results from another M&S are also used for V&V.
4. Significant				
No data is available for the real system in the real environment.	&	Ample test data for similar systems in similar environments is available.	&	No other Engineering Simulation data is used for V&V.
OR				
No data is available for the real system in the real environment.	&	No data is available for a similar system in similar environment.	&	Credible results from another M&S are also used for V&V.
5. Controlling				
No data is available for the real system in the real environment.	&	No data is available for a similar system in similar environment.	&	No other Engineering Simulation data is used for V&V.

A Usage Impact Rating Matrix may be developed by combining the Decision Consequence on one axis with the Engineering Simulation Influence on the other axis.

Usage Impact Rating Matrix

A Usage Impact Rating Matrix may be developed by combining the Decision Consequence on one axis with the Engineering Simulation Influence on the other axis. Rather than using the red / yellow / green approach for this matrix, calculation of a Usage Impact Rating is proposed using on the average of each ranking for a particular position in the matrix as illustrated in the table below.

Usage Impact Rating Matrix						
Engineering Simulation Influence Ranking	5	3	3.5	4	4.5	5
	4	2.5	3	3.5	4	4.5
	3	2	2.5	3	3.5	4
	2	1.5	2	2.5	3	3.5
	1	1	1.5	2	2.5	3
		1	2	3	4	5
		Decision Consequence Ranking				

The Usage Impact Rating values from the matrix help to understand the impact of the expected usage of the Engineering Simulation.

Determination of Applicable Credibility Reviews by Phase

This paper explores potential Credibility Reviews for each phase and an initial concept for attributes to be considered for each review. Each Credibility Review involves determining the ranking score for that criteria and comparing that to the associated Credibility Objective for that review.

Credibility Objectives

NASA-STD-7009A refers to Credibility Assessment Thresholds while the PCMM refers to Maturity Level Requirements. The generalized ESRM outlined in this paper uses the term Credibility Objectives for establishing the desired scoring of the Credibility Reviews to support confidence in the current Engineering Simulation activity as it relates to each Engineering Simulation phase.

One approach to establishing Credibility Objectives is to have a team that consists of all the appropriate members which determines the objective values for each applicable Credibility Review for the Engineering Simulation under consideration. This approach is consistent with the recommendations of NASA-STD-7009A and PCMM.

Another approach to establishing Credibility Objectives is to use a single common base objective value. One possible method for this approach is to establish a Usage Impact Credibility Objective value based on the Usage Impact Rating Matrix. The benefit of this approach is that it relates the Credibility Objectives to the criticality associated with the decision being influenced by the Engineering Simulation activity as it relates to each Engineering Simulation phase.

The following table illustrates the resulting Usage Impact Credibility Objectives.

Usage Impact Credibility Objectives						
Engineering Simulation Influence Ranking	5	3	3.5	4	4.5	5
	4	2.5	3	3.5	4	4.5
	3	2	2.5	3	3.5	4
	2	1.5	2	2.5	3	3.5
	1	1	1.5	2	2.5	3
		1	2	3	4	5
		Decision Consequence Ranking				

One approach to establishing Criteria Objectives is to have a team that consists of all the appropriate members which determines the objective values for each applicable criterion for the Engineering Simulation phase under consideration.

Any common base objective value approach to Credibility Objectives needs to account for the fact that every criterion being assessed does not have the same importance to the overall assessment of risk and credibility, and the values probably may need to be adjusted by an applicability factor for each criterion. The concerns related to this approach are:

1. It is an unverified quantification of qualitative assessments
2. Use of an applicability factor has the potential of abstracting information and providing a means to "game" the system.

The Credibility Objectives targeted for each Credibility Review should be determined prior to the start of any Engineering Simulation activity. The objective values may then be adjusted based on feasibility of achieving the objectives which in turn reduce (or increase) the domain of applicability.

Another approach to establishing Criteria Objectives is to use a single common base objective value. One possible method for this approach is to establish a common base objective value based on the Usage Impact Rating Matrix.

Credibility Reviews

It is proposed that the applicable Credibility Reviews for each Engineering Simulation phase should be assessed using a range of values from 1 to 5 but not limited to integer values. Each Credibility Review rating will have multiple attributes to be evaluated. It is proposed that a non-integer Credibility Review rating be determined by averaging all the appropriate attribute evaluations. The use of non-integer values for Credibility Review rating obviates the need for defining detailed sub-criteria.

It is also recommended that the minimum to maximum range be evaluated. For example, a criterion having 3 attributes rated at level 1, 1 at level 2, and 1 at level 3 would have a criterion assessment rating of 1.6, with a minimum rating of 1 and a maximum rating of 3. It is not recommended to try to determine an overall Credibility Review rating. The overall evaluation should instead be part of the Appropriateness Assessment.

The detailed Credibility Review attribute-based ratings are different according to the specific phase of Engineering Simulation Activity, where the definition of the attribute-based ratings for each phase is beyond the scope of this paper.

Appropriateness Assessment

An Appropriateness Assessment can be performed for each phase of Engineering Simulation based on the Credibility Objectives and Credibility Reviews proposed earlier in this paper. The purpose of the Appropriateness Assessment is to determine if the Engineering Simulation activity that was performed is "appropriate" for its intended use, including the expected influence it will have on the technical or business decision being made.

One possible approach for the Appropriateness Assessment is based on a quantifiable “Appropriateness Index.” An Appropriateness Index can be calculated for each Credibility Review by dividing each applicable Credibility Review Rating by the corresponding Credibility Objective value. The Appropriateness Indices for a simulation under review should be calculated for the minimum, mean and maximum Credibility Review Ratings. A set of overall Appropriateness Indices can then be calculated by averaging the Appropriateness Indices for min, mean, and average Appropriateness Indices of each applicable Credibility Review.

The purpose of the Appropriateness Assessment is to determine if the simulation that was performed is “appropriate” for the expected influence it will make on the decision being made.

An Appropriateness Index of less than 1.0 indicates that the simulation performed is not appropriate to support the intended use as related to the decision being made. This approach results in multiple appropriateness qualifications as follows:

- Clearly appropriate
    - Minimum Appropriateness Index is equal to or greater than 1.0
- Possibly appropriate and needs further investigation
    - Mean Appropriateness Index is equal to or greater than 1.0
    - Minimum Appropriateness Index is less than 1.0
- Clearly not appropriate
    - Mean Appropriateness Index is less than 1.0

A corresponding “Risk Index” related to the anticipated decision support by a specific simulation can also be calculated by subtracting the Appropriateness Index from 1.0 with an adjustment that any negative risk indices are set to 0.0. When the Risk Index is greater than 0.0, there is risk associated with the simulation being used to support the decision under consideration.

One possible additional consideration to be explored is the feasibility that a Usage Impact Appropriateness Matrix could be determined based on the appropriateness indices calculated for the full matrix of Usage Impact Credibility Objectives. This would provide an understanding of the “Domain of Appropriateness” for the software, methodologies & process, and application of Engineering Simulations to the impact of the decision to be made. This can offer guidance on the scope of use (Algorithms and Software, Methodology & Process) and reuse (Application).

A Usage Impact Appropriateness Matrix could be developed for each Engineering Simulation activity through the following process:

- For each Applicable Credibility Review
  - Determine an Applicability Factor
- For each cell in the Usage Impact Credibility Objective Matrix
  - Set the common base objective to the cell value
  - For each Credibility Review
    - Multiply the Applicability Factor by the Usage Impact Credibility Objective to determine the appropriate Credibility Objective
    - Determine the mean and minimum Appropriateness Indices
  - Aggregate a composite mean and minimum Appropriateness Index by averaging the mean and minimum indices for all Credibility Reviews
  - If minimum aggregate Appropriateness is equal to or greater than 1.0
    - Mark as "Yes"
  - If mean aggregate Appropriateness is equal to or greater than 1.0 and minimum aggregate Appropriateness is less than 1.0
    - Mark as "Maybe"
  - If mean aggregate Appropriateness Index is less than 1.0
    - Mark as "No"

One possible additional consideration to be explored is the feasibility that a Usage Impact Appropriateness Matrix could be determined based on the appropriateness indices calculated for the full matrix of Usage Impact Credibility Objectives.

The following table illustrates what an Impact Based Appropriateness Matrix would look like.

Usage Impact Appropriateness Matrix						
Engineering Simulation Influence Ranking	5	Maybe	Maybe	No	No	No
	4	Maybe	Maybe	Maybe	No	No
	3	Yes	Maybe	Maybe	Maybe	No
	2	Yes	Yes	Maybe	Maybe	Maybe
	1	Yes	Yes	Yes	Yes	Maybe
		1	2	3	4	5
		Decision Consequence Ranking				

Phase-Based Predictive Capability Assessment

Earlier in this paper, it was mentioned that Engineering Simulation activities can be grouped into three major phases:

1. Algorithm & Software Development
2. Methodology & Process Development
3. Methodology & Process Application

This separation into three distinct phases enables a generalized approach that separates the Predictive Capability Assessment of each phase. The proposed Credibility Reviews for each Engineering Simulation phase will include some (but not all) of the following as is appropriate for that phase:

- Previous Phase (if applicable)
- Usage
- Pedigree
- Verification
- Fidelity
- Validation
- Uncertainty
- Robustness

Phase 1: Algorithm & Software Development

The Algorithm & Software Development phase described in this paper can be performed by commercial software companies or internal software development teams. NASA and Sandia National Laboratories both do significant simulation algorithm and software development and have appropriately made this a strong emphasis of their respective assessments. Commercial companies may do some internal software development; however, the vast majority of analysis is performed via commercial simulation software.

Commercial software vendors are highly likely to have their own QA processes, including suitability of use that are much more sophisticated and comprehensive than what we could explore in this paper. However, some commercial software vendors may consider portions of this process including pedigree, verification, and validation as intellectual property. NAFEMS, over the years, has developed a neutral set of benchmarks for a range of physics applications that can be used as a resource for software verification of commercial simulation software.

The discussions of the Algorithm & Software Development phase in this paper are limited to internal development which should include "user defined" materials or solution algorithms linked to commercial software.

The Algorithm & Software Development phase described in this paper can be performed by commercial software companies or internal software development teams.

Commercial companies may do some internal software development; however, the vast majority of analysis is performed via commercial simulation software.

Phase 1 Applicable Credibility Reviews

For those who are engaged in internal simulation software development, the proposed Applicable Credibility Reviews for the Algorithm & Software Development phase are:

- Usage
- Data Pedigree
- Verification
- Fidelity
- Robustness

Phase 1 Credibility Objectives

The Credibility Objectives for each Credibility Review should be determined prior to the start of any activity in the Algorithm & Software Development Phase based on the desired Domain of Applicability. The objective values may then be adjusted based on feasibility of achieving the Credibility Objectives resulting in a revised Domain of Applicability.

When considering "appropriateness" for Phase 1, it may be desirable to establish a Domain of Appropriateness rather than Appropriateness Indices for a specific proposed usage. One advantage of the Usage Impact Credibility Objectives approach is that a Usage Impact Algorithm & Software Appropriateness Matrix could be determined to define the Domain of Appropriateness. This Domain of Appropriateness could then be used by subsequent phases of Engineering Simulation activity to determine potential "appropriateness" of the underlying algorithm and software prior to execution of that phase.

When considering "appropriateness" for Phase 1 it may be desirable to establish a Domain of Appropriateness rather than Appropriateness Indices for a specific proposed usage.

Usage Impact
Algorithm & Software Appropriateness Matrix

Engineering Simulation Influence Ranking	5	Maybe	Maybe	No	No	No
	4	Maybe	Maybe	Maybe	No	No
	3	Yes	Maybe	Maybe	Maybe	No
	2	Yes	Yes	Maybe	Maybe	Maybe
	1	Yes	Yes	Yes	Yes	Maybe
		1	2	3	4	5
		Decision Consequence Ranking				

Phase 1 Credibility Reviews

The detailed Credibility Reviews and scoring in this phase are beyond the scope of this paper; however, for each Credibility Review, a first pass at items to be considered in this phase are outlined below.

Each of the items listed below may have several attributes to be reviewed for each level. The proposed approach is to average the scoring for all attributes as a result of any specific Credibility Review, resulting in a non-integer scoring with a min, mean, and max value.

Phase 1 Credibility Review: Usage

The following table outlines a first pass at items to be considered during the Credibility Review for Usage in this phase.

Usage	
	Intended use
	Permissible uses
	Domain of Applicability
	Application needs
	Assumptions & limitations
	Quantities of Interest
	Usage Credibility Review compliance

Phase 1 Credibility Review: Pedigree

The following table outlines a first pass at items to be considered during the Credibility Review for Pedigree in this phase.

Data Pedigree	
	Reference data
	Verification data
	Pedigree Credibility Review compliance

Phase 1 Credibility Review: Verification Review

The following table outlines a first pass at items to be considered during the Credibility Review for Verification in this phase.

Verification	
	SQE process compliance
	Test Coverage
	Attributes & errors
	Algorithm & Software verification
	Assumptions & limitations
	Domain of Verification
	Verification Credibility Review compliance

Phase 1 Credibility Review: Fidelity

The Fidelity review in this phase is mainly focused on clear explanation of expected, acceptable, and required data and model fidelity for the permissible uses and the resulting domain of applicability. The following table outlines a first pass at items to be considered during the Credibility Review for Fidelity in this phase.

The detailed Credibility Reviews and scoring in this phase are beyond the scope of this paper; however, for each Credibility Review a first pass at items to be considered in this phase are outlined.

Fidelity	
	Physics fidelity
	Representation fidelity
	Material fidelity
	Initial conditions
	Boundary conditions
	Fidelity Credibility Review compliance

Phase 1 Credibility Review: Robustness

The following table outlines a first pass at items to be considered during the Credibility Review for Robustness in this phase.

Robustness	
	Sensitivity to data
	Sensitivity to fidelity
	Robustness Credibility Review compliance

Phase 2: Methodology & Process Development

The second phase of Methodology & Process Development as described in this paper is typically performed by simulation experts who are becoming less likely to be the target user. Assessment of this phase is key, and it could be argued that the scoring in the third phase should not exceed the scoring in this phase for matching Credibility Reviews.

It is recommended that "ad-hoc" simulations that are not based on a previously-defined methodology/process should be considered initially in Phase 2 to assess the "appropriateness" of the approaches being used. The alternative would be to consider them as Phase 3 activities but with the Methodology & Process assessment status review set to 1.0, thereby reducing the potential "appropriateness."

The new emerging area of Simulation Applications (SimApps) is taking this phase into a new level that also includes embedded models. Additional considerations need to be given to Credibility Reviews when associated with SimApps.

It is recommended that "ad-hoc" simulations that are not based on a previously-defined methodology/ process should be considered initially in Phase 2 to assess the "appropriateness" of the approaches being used.

Phase 2 Applicable Credibility Reviews

The proposed Applicable Credibility Reviews for the Methodology & Process Development Phase are:

- Algorithm & Software
- Usage
- Data & History Pedigree
- Verification
- Fidelity
- Validation
- Uncertainty
- Robustness

Phase 2 Credibility Objectives

The Credibility Objectives for each Credibility Review should be determined prior to the start of any activity in the Methodology & Process Development Phase based on the desired Domain of Applicability. The objective values may then be adjusted based on feasibility of achieving the objectives and possibly resulting in a revised Domain of Applicability.

When considering "appropriateness" for this phase it may be desirable to establish a Domain of Appropriateness in addition to Appropriateness Indices for a specific proposed usage. A Usage Impact Methodology & Process Appropriateness Matrix could be determined to define the Domain of Appropriateness for the specific methodology or process being evaluated. This Domain of Appropriateness could then be used during the application phase to determine potential "appropriateness" of the underlying methods and processes prior to their application.

A Methodology & Process Appropriateness Matrix could be determined to define the Domain of Appropriateness for the specific methodology or process being evaluated.

Usage Impact
Methodology & Process Appropriateness Matrix

Engineering Simulation Influence Ranking	5	Maybe	Maybe	No	No	No
	4	Maybe	Maybe	Maybe	No	No
	3	Yes	Maybe	Maybe	Maybe	No
	2	Yes	Yes	Maybe	Maybe	Maybe
	1	Yes	Yes	Yes	Yes	Maybe
		1	2	3	4	5
		Decision Consequence Ranking				

Phase 2 Credibility Reviews

The detailed Credibility Reviews and scoring in this phase are beyond the scope of this paper; however, for each Credibility Review a first pass at items to be considered in this phase are outlined below.

Each of the items listed may have several attributes to be reviewed for each ranking level. The proposed approach is to average the scoring for all attributes in this Credibility Review resulting in a non-integer scoring with a min, mean, and max value.

Each of the items listed may have one or several attributes to be reviewed for each ranking level. The proposed approach is to average the scoring for all attributes in this criterion resulting in a non-integer scoring with a min, mean, and max value.

Phase 2 Credibility Review: Algorithm & Software

This Credibility Review is slightly different from the other Credibility Reviews in that it incorporates the status of the previous phase. One approach is to include this review of the previous phase with an equal weighting to the other assessment reviews and thereby influencing the aggregate mean and minimum Appropriateness Indices. There is also a distinct possibility that the previous phase assessment has not been completed. The following table outlines a first pass at items to be considered during the Credibility Review for Algorithm & Software in this phase.

Algorithm & Software	
	Algorithm & Software assessment ratings
	Robustness scores for similar Methods & Processes
	Validation scores for similar Methods & Processes

Phase 2 Credibility Review: Usage

The following table outlines a first pass at items to be considered during the Credibility Review for Usage in this phase.

Usage	
	Intended use
	Permissible uses
	Domain of Applicability
	Application needs
	Assumptions & limitations
	Quantities of interest
	Usage Credibility Review compliance

Phase 2 Credibility Review: Data & History Pedigree

The following table outlines a first pass at items to be considered during the Credibility Review of Data & History Pedigree in this phase.

Data & History Pedigree	
	Reference data
	Verification data
	Validation data
	Input data
	History
	Management
	Pedigree Credibility Review compliance

Phase 2 Credibility Review: Verification

The following table outlines a first pass at items to be considered during the Credibility Review of Verification in this phase.

Verification	
	Test coverage
	Attributes & errors
	Method & Process verification
	Solution verification
	Assumptions & limitations
	Domain of Verification
	Verification Credibility Review compliance

Phase 2 Credibility Review: Fidelity

The following table outlines a first pass at items to be considered during the Credibility Review of Fidelity in this phase.

Fidelity	
	Physics fidelity
	Representation fidelity
	Material fidelity
	Initial conditions
	Boundary conditions
	Fidelity Credibility Review compliance

Phase 2 Credibility Review: Validation

The following table outlines a first pass at items to be considered during the Credibility Review of Validation in this phase.

Validation	
	Problem statement
	Validation criteria
	Accuracy requirements
	Validation Experiment definition
	Domain of Validation
	Validation Credibility Review compliance

The proposed approach is to average the scoring for all attributes in this criterion resulting in a non-integer scoring with a min, mean, and max value.

Phase 2 Credibility Review: Uncertainty

The following table outlines a first pass at items to be considered during the Credibility Review of Uncertainty in this phase.

Uncertainty	
	Uncertainty identification
	Uncertainty quantification
	Impact on Quantities of Interest
	Uncertainty Credibility Review compliance

Phase 2 Credibility Review: Robustness

The following table outlines a first pass at items to be considered during the Credibility Review of Robustness in this phase.

Robustness	
	Results extraction methodology
	Sensitivity to data
	Sensitivity to fidelity
	Sensitivity to uncertainty
	Robustness Credibility Review compliance

Phase 3: Methodology & Process Application

The third phase of Methodology & Process Application as described in this paper is typically performed by a wide range of Engineering Simulation users from simulation experts to simulation novices. Assessment of Phase 3 activity is key for determining the "appropriateness" of any given Engineering Simulation to support the technical and/or business decision at hand, and the risk associated with the Engineering Simulation Influence on that decision.

Applicable Credibility Reviews

The proposed Applicable Credibility Reviews for the Methodology & Process Application phase are:

- Methodology & Process Development
- Usage
- Data & History Pedigree
- Verification
- Fidelity
- Validation
- Uncertainty
- Robustness

Assessment of Phase 3 activity is key for determining the "appropriateness" of any given Engineering Simulation to support the technical and/or business decision at hand, and the risk associated with the Engineering Simulation Influence on that decision.

Credibility Objectives

The Credibility Objectives for each Credibility Review should be determined prior to the start of the Methodology & Process Application activity based on the Proposed Usage.

A Usage Impact Application Appropriateness Matrix could also be determined to define the Domain of Appropriateness to determine potential "appropriateness" for reuse.

Usage Impact Application Appropriateness Matrix						
Engineering Simulation Influence Ranking	5	Maybe	Maybe	No	No	No
	4	Maybe	Maybe	Maybe	No	No
	3	Yes	Maybe	Maybe	Maybe	No
	2	Yes	Yes	Maybe	Maybe	Maybe
	1	Yes	Yes	Yes	Yes	Maybe
		1	2	3	4	5
		Decision Consequence Ranking				

The Criteria Objectives for each Credibility Review should be determined prior to the start of the Methodology & Process Application activity based on the Proposed Usage.

Phase 3 Credibility Review: Methodology & Process Development

This Credibility Review incorporates an assessment of the previous phase of activity. The following table outlines a first pass at items to be considered during the Credibility Review of Methodology & Process in this phase.

Methodology & Process Development	
	Methodology & Process assessment ratings
	Robustness Scores for similar applications
	Validation Scores for similar applications
	Uncertainty Scores for similar applications

Phase 3 Credibility Review: Usage

The following table outlines a first pass at items to be considered during the Credibility Review of Usage in this phase.

Usage	
	Permissible uses
	Domain of applicability
	Application needs
	Assumptions & limitations
	Quantities of interest
	Usage Credibility Review compliance

Phase 3 Credibility Review: Data & History Pedigree

The following table outlines a first pass at items to be considered during the Credibility Review of Data & History Pedigree in this phase.

Data & History Pedigree	
	Input data
	History
	Management
	Pedigree Credibility Review compliance

Phase 3 Credibility Review: Verification

The following table outlines a first pass at items to be considered during the Credibility Review of Verification in this phase.

Verification	
	Attributes & errors
	Method & Process verification
	Solution verification
	Assumptions & limitations
	Domain of Verification
	Verification Credibility Review compliance

Phase 3 Credibility Review: Fidelity

The following table outlines a first pass at items to be considered during the Credibility Review of Fidelity in this phase.

Fidelity	
	Physics fidelity
	Representation fidelity
	Material fidelity
	Initial conditions
	Boundary conditions
	Fidelity Credibility Review compliance

The third phase of Methodology & Process Application as described in this paper is typically performed by a wide range of Engineering Simulation users from simulation experts to simulation novices.

Phase 3 Credibility Review: Validation

The following table outlines a first pass at items to be considered during the Credibility Review of Validation in this phase.

Validation	
	Problem statement
	Validation criteria
	Accuracy requirements
	Validation Experiment definition
	Domain of Validation
	Validation Credibility Review compliance

Phase 3 Credibility Review: Uncertainty

The following table outlines a first pass at items to be considered during the Credibility Review of Uncertainty in this phase.

Uncertainty	
	Uncertainty identification
	Uncertainty quantification
	Impact on Quantities of Interest
	Uncertainty Credibility Review compliance

Phase 3 Credibility Review: Robustness

The following table outlines a first pass at items to be considered during the Credibility Review of Robustness in this phase.

Robustness	
	Results extraction methodology
	Sensitivity to data
	Sensitivity to fidelity
	Sensitivity to uncertainty
	Robustness Credibility Review compliance

Sample Illustration

To illustrate how a generalized ESRM might be used, a sample illustration based on an artificial scenario will be presented. We will not be focusing on the detailed attributes that led to ranking or scoring in each area but instead will illustrate how these can be used to understand “appropriateness” and risk.

The artificial scenario presented in this paper is an assessment of a finite element analysis performed by a new analyst based on a well-known and documented standard work process. The assessment is being performed by the engineering manager.

The artificial scenario presented in this paper is an assessment of a finite element analysis performed by a new analyst based on a well-known and documented standard work process.

The first step as recommended in this paper is for the engineering manager to determine a Usage Impact Rating. This involves understanding the Decision Consequence and the Engineering Simulation Influence.

In our illustrative scenario, the engineering manager has assessed the Decision Consequence as Minor (Level 2) based on the rankings proposed in this paper. The engineering manager has also assessed the Engineering Simulation Influence to be at Moderate (Level 3). Based on these assessments, the engineering manager can now determine the Usage Based Impact Factor resulting in a Usage Impact Rating of 2.5.

Usage Impact Rating Matrix						
Engineering Simulation Influence Ranking	5					
	4					
	3		2.5			
	2					
	1					
		1	2	3	4	5
		Decision Consequence Ranking				

For the purpose of illustration, our engineering manager elected to define Credibility Objectives using the approach of a common base value with Applicability Factors. Our engineering manager also elected to use the Usage Impact Rating that was determined for this analysis as the common base value for Credibility Objectives (Usage Impact Credibility Objectives).

This scenario clearly falls into Phase 3: Methodology & Process Application, where the engineering manager used the Credibility Reviews outlined in this paper for Phase 3.

The following table illustrates the determination of Credibility Objectives by the engineering manager for this sample scenario. Since the analysis is based on a well-known and documented standard work process, the engineering manager has chosen to reduce the Applicability Factors for Methodology & Process Review, Verification Review, and Uncertainty Review.

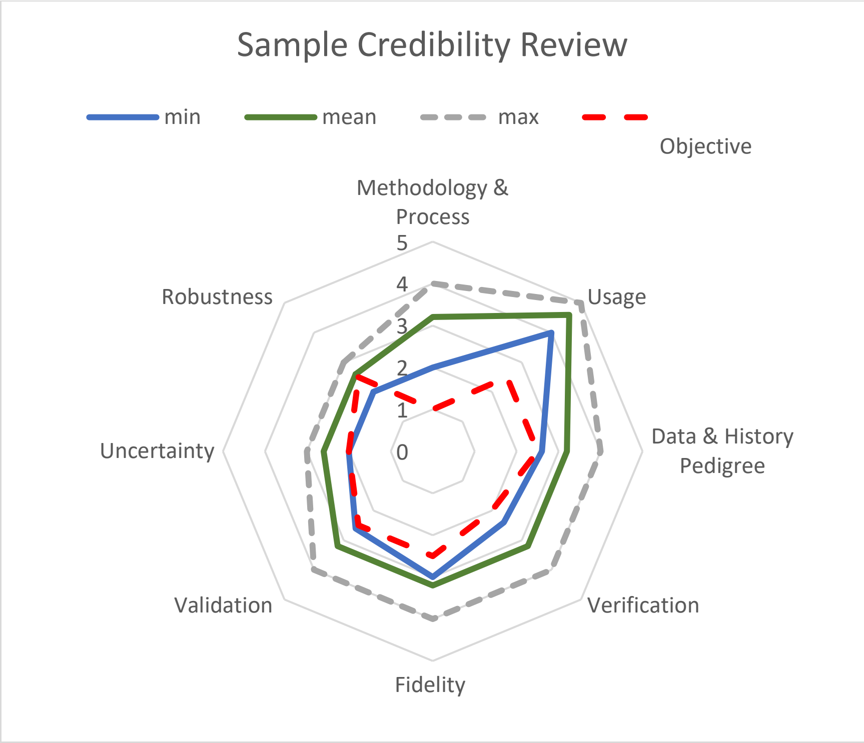
This scenario clearly falls into the Methodology & Process Application phase (Phase 3) and we will use the Criteria outlined for that phase.

Phase 3 Reviews	Applicability Factor	base objective	Credibility Objectives
Methodology & Process	0.4	2.5	1
Usage	1	2.5	2.5
Data & History Pedigree	1	2.5	2.5
Verification	0.8	2.5	2
Fidelity	1	2.5	2.5
Validation	1	2.5	2.5
Uncertainty	0.8	2.5	2
Robustness	1	2.5	2.5

The next step in our illustrative sample is for the engineering manager and analysts to perform a Credibility Review for each of the applicable criteria. The following table and radar plot illustrate the sample findings of this Credibility Review.

	Credibility Review		
Phase 3 Reviews	min	mean	max
Methodology & Process	2	3.2	4
Usage	4	4.6	5
Data & History Pedigree	2.6	3.2	4
Verification	2.4	3.2	4
Fidelity	3	3.2	4
Validation	2.6	3.2	4
Uncertainty	2	2.6	3
Robustness	2	2.6	3

The next step in our illustrative sample is for the engineering manager and analysts to perform a Credibility Assessment for each of the applicable criteria.



Based on the mean values of the Credibility Reviews this analysis appears to meet the Credibility Objectives for all applicable criteria.

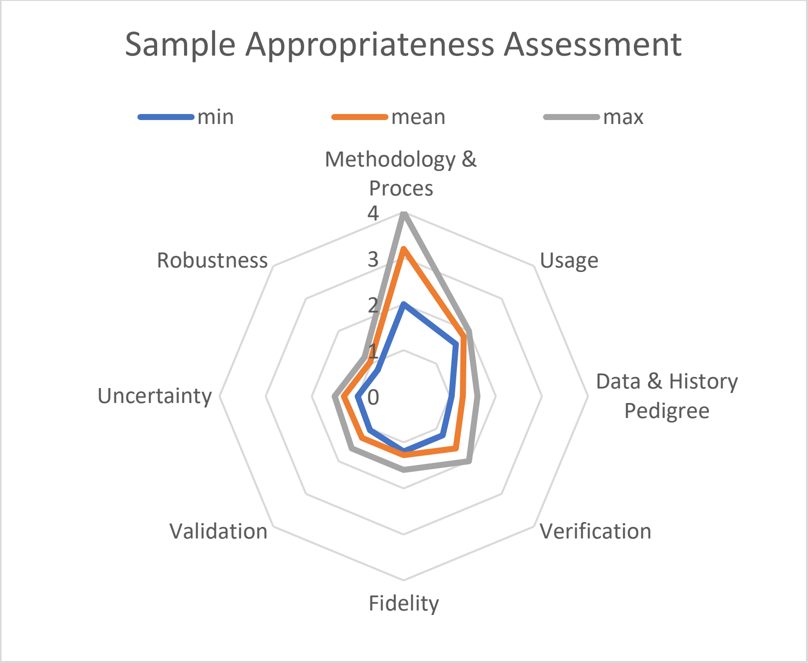
Even with this promising set of data, the engineering manager in our scenario has chosen to perform an Appropriateness Assessment as proposed earlier in this paper.

The minimum, mean, and maximum Appropriateness Indices are calculated for each criterion by dividing the assessment value by the objective value for each criterion. An aggregate set of Appropriateness Indices is also calculated as the mean value of the minimum, mean, and maximum Appropriateness Index Values.

The following table and radar plot illustrate the Appropriateness Indices for the sample scenario.

Based on the mean values of assessment that this analysis appears to meet the Credibility Objectives for all applicable criteria.

	Appropriateness Indices		
Phase 3 Reviews	min	mean	max
Methodology & Process	2	3.2	4
Usage	1.6	1.84	2
Data & History Pedigree	1.04	1.28	1.6
Verification	1.2	1.6	2
Fidelity	1.2	1.28	1.6
Validation	1.04	1.28	1.6
Uncertainty	1	1.3	1.5
Robustness	0.8	1.04	1.2
Aggregate Appropriateness Index	1.235	1.6025	1.9375



Since the aggregate minimum appropriateness index is greater than 1.0, the engineering manager classified the analysis as "Clearly Appropriate." However, our engineering manager is rather thorough and proceeded to calculate corresponding Risk Indices as outlined in this paper.

The following table illustrates that even though the analysis is clearly appropriate using this analysis as planned for the decision at hand, it is not without risk. There is a small risk related to the Robustness Review of the simulation. The engineering manager may choose one of the following: 1. Accept the small risk, 2. Review the Applicability Factor used, or 3. Ask the analyst to address the issue.

Even with this promising set of data, the engineering manager in our scenario has chosen to perform an Appropriateness Assessment as proposed earlier in this paper.

	Risk Indices		
Phase 3 Reviews	min	mean	max
Methodology & Process	0	0	0
Usage	0	0	0
Pedigree	0	0	0
Verification	0	0	0
Fidelity	0	0	0
Validation	0	0	0
Uncertainty	0	0	0
Robustness	0	0	0.2
Aggregate Risk Index	0	0	0.025

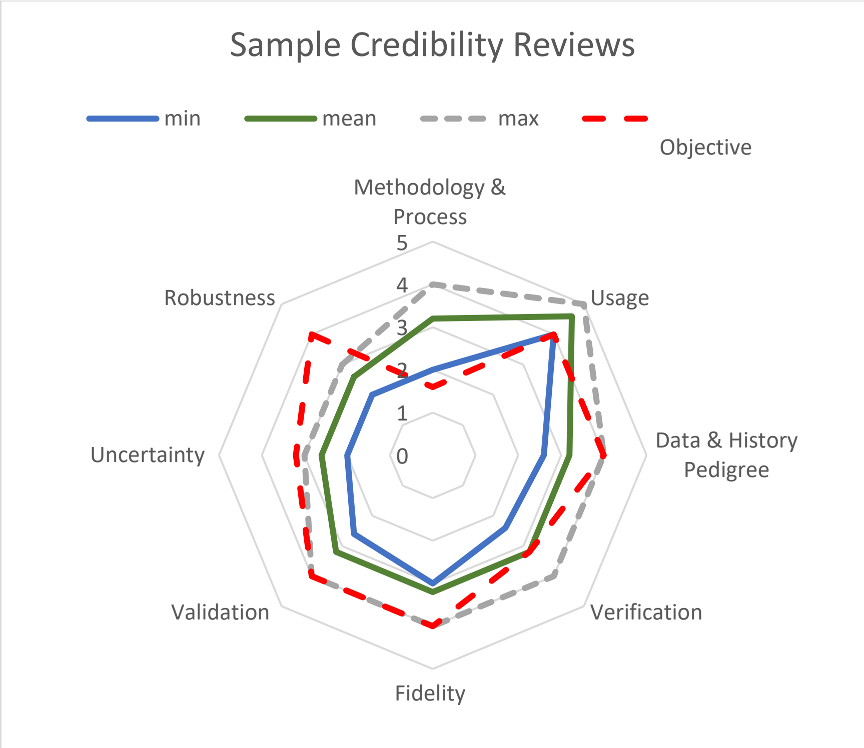
By now the curiosity of our engineering manager is piqued and wants to understand better where this analysis might be reused to support other decisions. The engineering manager decides to calculate a Domain of Appropriateness for this analysis using the full range of Usage Impact ratings as objectives.

Usage Impact Domain of Appropriateness						
Engineering Simulation Influence Ranking	5	Yes	Maybe	Maybe	No	No
	4	Yes	Yes	Maybe	Maybe	No
	3	Yes	Yes	Yes	Maybe	Maybe
	2	Yes	Yes	Yes	Yes	Maybe
	1	Yes	Yes	Yes	Yes	Yes
		1	2	3	4	5
		Decision Consequence Ranking				

The engineering manager, being thorough, then looked at the radar plot for one of the “Maybe” use cases for a clear understanding of what would need to be improved in this analysis for broader reuse. Based on the radar plot, the engineering manager can see that to meet this broader target usage, the following Credibility Review ratings would have to be improved:

- Data & History Pedigree
- Fidelity
- Validation
- Uncertainty
- Robustness

There is a small risk related to the Robustness Review of the simulation. The engineering manager may choose one of the following: 1. Accept the small risk, 2. Review the Applicability Factor used, or 3. Ask the analyst to address the issue.



The radar chart also shows that the Credibility Review indicated that the biggest focus for required improvement should be on Robustness.

SUMMARY

This ASSESS Strategic Insight Paper provides a quick review of previous efforts related to M&S assessments at NASA (NASA-STD-7009A) and Sandia National Laboratories (Predictive Capability Maturity Model), along with a proposed generalized Predictive Capability Assessment approach based on a generalized ESRM.

The proposed generalized Engineering Simulation Risk Model (ESRM) provides a set of recommendations, reviews, and criteria that could be used in support of Engineering Simulation influenced decisions. The proposed ESRM in this paper leverages information and principles from NASA-STD-7009A and Sandia PCMM.

The proposed approach provides guidelines and a potential framework for defining an ESRM based on evaluation of three separate Engineering Simulation phases:

1. Algorithm & Software Development
2. Methodology & Process Development
3. Methodology & Process Application

The proposed generalized Engineering Simulation Risk Model (ESRM) provides a set of recommendations and criteria that could be used in support of Engineering Simulation influenced decisions.

The generalized ESRM outlined in this paper is a Predictive Capability Assessment that consists of:

a) Usage Impact
b) Phase-Based Predictive Capability Assessment
    a. Determination of Applicable Credibility Reviews
    b. Credibility Objectives
    c. Credibility Reviews
    d. Appropriateness Assessment

The primary purpose of the proposed generalized ESRM is to improve understanding of the "credibility" of a simulation, thereby increasing confidence in the Engineering Simulation influenced decisions. The secondary purpose is to initiate discussions leading to a potential path to a scalable ESRM that can be broadly deployed.

The objective of this ASSESS Strategic Insight Paper is to initiate discussions that can lead to further development of an effective generalized Engineering Simulation Risk Model that can be used to support the current and dramatically expanding use of Engineering Simulation. Understanding and enabling an effective and consistent ESRM are key elements that are required for a significantly broader use of Engineering Simulation in support of more informed business and technical decision making.

The objective of this ASSESS Strategic Insight Paper is to initiate discussions that can lead to further development of an effective generalized Engineering Simulation Risk Model that can be used to support the current and dramatically expanding use of Engineering Simulation.

REFERENCES

1. NASA-STD-7009A w/ CHANGE 1: ADMINISTRATIVE/EDITORIAL CHANGES - https://standards.nasa.gov/standard/nasa/nasa-std-7009
2. Predictive Capability Maturity Model for Computational Modeling and Simulation - SANDIA REPORT SAND2007-5948 - https://cfwebprod.sandia.gov/cfdocs/CompResearch/docs/Oberkampf-Pilch-Trucano-SAND2007-5948.pdf
3. Predictive Capability in Computational Science and Engineering - 2007 Salishan Conference Confidence in Predictive Simulations (SAND2007-2316C) - https://www.lanl.gov/conferences/salishan/salishan2007/Salishan-07-4-final.W.Oberkampf.pdf

