NORTH ATLANTIC TREATY ORGANIZATION

SCIENCE AND TECHNOLOGY ORGANIZATION

AC/323(MSG-139)TP/945

www.sto.nato.int

STO TECHNICAL REPORT **TR-MSG-139**

Guidelines for Modelling and Simulation (M&S) Use Risk Identification, Analysis, and Mitigation

(Lignes directrices d'identification, d'analyse et d'atténuation du risque d'utilisation de la modélisation et simulation (M&S))

Final Report of Modelling and Simulation Group MSG-139.

Published September 2021

Distribution and Availability on Back Cover

NORTH ATLANTIC TREATY ORGANIZATION

SCIENCE AND TECHNOLOGY ORGANIZATION

AC/323(MSG-139)TP/945

www.sto.nato.int

STO TECHNICAL REPORT **TR-MSG-139**

Guidelines for Modelling and Simulation (M&S) Use Risk Identification, Analysis, and Mitigation

(Lignes directrices d'identification, d'analyse et d'atténuation du risque d'utilisation de la modélisation et simulation (M&S))

Final Report of Modelling and Simulation Group MSG-139.

The NATO Science and Technology Organization

Science & Technology (S&T) in the NATO context is defined as the selective and rigorous generation and application of state-of-the-art, validated knowledge for defence and security purposes. S&T activities embrace scientific research, technology development, transition, application and field-testing, experimentation and a range of related scientific activities that include systems engineering, operational research and analysis, synthesis, integration and validation of knowledge derived through the scientific method.

In NATO, S&T is addressed using different business models, namely a collaborative business model where NATO provides a forum where NATO Nations and partner Nations elect to use their national resources to define, conduct and promote cooperative research and information exchange, and secondly an in-house delivery business model where S&T activities are conducted in a NATO dedicated executive body, having its own personnel, capabilities and infrastructure.

The mission of the NATO Science & Technology Organization (STO) is to help position the Nations' and NATO's S&T investments as a strategic enabler of the knowledge and technology advantage for the defence and security posture of NATO Nations and partner Nations, by conducting and promoting S&T activities that augment and leverage the capabilities and programmes of the Alliance, of the NATO Nations and the partner Nations, in support of NATO's objectives, and contributing to NATO's ability to enable and influence security and defence related capability development and threat mitigation in NATO Nations and partner Nations, in accordance with NATO policies.

The total spectrum of this collaborative effort is addressed by six Technical Panels who manage a wide range of scientific research activities, a Group specialising in modelling and simulation, plus a Committee dedicated to supporting the information management needs of the organization.

- AVT Applied Vehicle Technology Panel
- HFM Human Factors and Medicine Panel
- IST Information Systems Technology Panel
- NMSG NATO Modelling and Simulation Group
- SAS System Analysis and Studies Panel
- SCI Systems Concepts and Integration Panel
- SET Sensors and Electronics Technology Panel

These Panels and Group are the power-house of the collaborative model and are made up of national representatives as well as recognised world-class scientists, engineers and information specialists. In addition to providing critical technical oversight, they also provide a communication link to military users and other NATO bodies.

The scientific and technological work is carried out by Technical Teams, created under one or more of these eight bodies, for specific research activities which have a defined duration. These research activities can take a variety of forms, including Task Groups, Workshops, Symposia, Specialists' Meetings, Lecture Series and Technical Courses.

The content of this publication has been reproduced directly from material supplied by STO or the authors.

Published September 2021

ISBN 978-92-837-2269-4

Table of Contents

List of Figures

List of Tables

List of Acronyms

CSO	Collaborative Support Office
DLR	Detailed Level Requirement
DoD	Department of Defense
GM-VV	Generic Methodology for Verification and Validation
HLR	High Level Requirement
IEEE	Institute of Electrical and Electronics Engineers
ITIS	Institut für Technik Intelligenter Systeme
JHU/APL	Johns Hopkins University Applied Physics Laboratory
M&S	Modelling and Simulation
MSG	Modelling and Simulation Group
MURM	M&S Use Risk Methodology
NATO	North Atlantic Treaty Organization
ONERA	Office National d'Etudes et Recherches Aérospatiales
QoI	Quantity of Interest
SISO	Simulation Interoperability Standards Office
SIU	Specific Intended Use
SNL	Sandia National Laboratory
STO	Science & Technology Organization
SW	Software
USC	University of Southern California
V&V	Verification and Validation
VV&A	Verification, Validation and Accreditationp

Glossary

Acceptance	The process that ascertains whether an M&S system is fit for specific intended use.
Acceptability Criteria	The criteria that the model, simulation, or federation of models and simulations needs to meet to be acceptable for its specific intended use.
Accreditation	The official certification that a model or simulation and its associated data are acceptable for use for a specific purpose.
Atoms	Decomposition of sub-factors.
Capability	What the model or simulation can do in terms of functional representations, behaviors, relationships, and interactions.
Factor	One of several areas that together characterize the state of risk for M&S requirements.
Information Entropy	The average amount of information produced by a stochastic source of data.
Information Theory	A theory that deals statistically with information, with the measurement of its content in terms of its distinguishing essential characteristics or by the number of alternatives from which it makes a choice possible, and with the efficiency of processes of communication between humans and machines.
Limitations	Restrictions in the ability of the M&S to represent the simuland with sufficient fidelity over the specific intended use problem space.
Referent	A codified body of knowledge about a thing being simulated.
Requirement	A singular documented physical or functional need that a particular design, product, or process aims to satisfy.
Simuland	The system being simulated by a simulation.
Sub-Factor	Decomposition of factors.
Tailoring	The modification of V&V processes, V&V organization, and V&V products to fit agreed risks, resources, and implementation constraints.
Validation	The process of determining the degree to which a model or simulation and its associated data are an accurate representation of the real world from the perspective of the specific intended uses of the model.
Verification	The process of determining that a model or simulation implementation and its associated data accurately represent the developer's conceptual description and specifications.

Acknowledgements

NMSG-139 would like to thank the following study observers:

- Ad van Lier, NL MoD JIVC/KIXS
- Manfred Rosa, Netherlands Aerospace Center, NLR
- Jeroen Voogd, Netherlands Organization for Applied Scientific Research (TNO)

NMSG-139 would also like to acknowledge the contributions made by others who were not officially part of the NMSG-139 Programme Committee:

- Lee Hendrix, report contributor;
- Peter Pandolfini, report contributor;
- Katherine Ruben, report coordinator; and
- Sebastian Sanchez, report contributor.

MSG-139 Membership List

CO-CHAIRS

Dr. Martin ADELANTADO
ONERA/CT/DTIM/MAGS
FRANCE
Email: Martin.Adelantado@onera.fr

Ms. Simone YOUNGBLOOD
Johns Hopkins University
UNITED STATES
Email: Simone.Youngblood@jhuapl.edu

MEMBERS

Lt. Col. Paolo DI BELLA
NATO Modelling & Simulation Centre of Excellence
ITALY
Email: dibellap@gmail.com

Lt. Col. Jan HODICKY
NATO Modelling & Simulation Centre of Excellence
CZECH REPUBLIC
Email: Jan.Hodicky@seznam.cz

Prof. Axel LEHMANN
Institut für Technik Intelligenter Systeme (ITIS)
GERMANY
Email: axel.lehmann@unibw.de

Mr. Giovanni Luca MAGLIONE
University of Genoa
ITALY
Email: gianluca.maglione@simulationteam.com

Capt. Vincenzo MILANO
NATO Modelling & Simulation Centre of Excellence
ITALY
Email: sesto.mes@smd.difesa.it

Mr. Dagfinn OLSON
Norwegian Defense Research Establishment (FFI)
NORWAY
Email: dagfinn.olson@ffi.no

ADDITIONAL CONTRIBUTORS

Mr. Grant BAILEY
UK MoD
UNITED KINGDOM
Email: Grant.Bailey127@mod.gov.uk

Major Tobias Uwe KUHN
NATO Modelling & Simulation Centre of Excellence
GERMANY
Email: Tobias-Uwe-Kuhn@web.de

Major Scott ROACH
Canadian Armed Forces
CANADA
Email: SCOTT.ROACH1@forces.gc.ca

Guidelines for Modelling and Simulation (M&S) Use Risk Identification, Analysis, and Mitigation

(STO-TR-MSG-139)

Executive Summary

In September 2014, the North Atlantic Treaty Organization (NATO) Collaboration Support Office (CSO) approved the formation of NATO Modelling and Simulation Group (NMSG) 139, "Modelling and Simulation (M&S) Use Risk Identification and Management" [1]. The major objective of this task group was to define and provide an initial implementation of a roadmap to support the development and deployment of a generic methodology, methods and techniques for M&S use risk identification, and analysis and to balance M&S use risk with resources applied to M&S verification and validation (V&V).

This task group proposed and formalized the M&S Use Risk Methodology (MURM) [2]. The task group analyzed the capability of MURM to balance risk with cost and to provide a means for risk identification and mitigation. In keeping with the rigorous approach utilized in the development of the MURM, MSG-139 initiated its work with a review of the related literature using the information identified and studied during the 2009 literature survey as the starting point [3]. The updated literature search confirmed that the MURM remains relevant to the problem space and that it is more mature than any other methodology. As such, the MURM was selected to be expanded and applied by MSG-139.

In the context of the MURM, M&S use risk is defined as [4]:

> *The probability that inappropriate application of M&S results for the specific intended use will produce unacceptable consequences to the decision maker.*

The approach taken in this work is to translate this definition into mathematical logic used to calculate M&S use risk on a requirement-by-requirement basis. The derivation of the methodology is based on coherent mathematical concepts that minimize unintended bias and establishes an explicit relationship to the V&V process and products. The MURM can be useful at several stages of the M&S development process.

This document reports the results of MSG-139 efforts including the literature survey but focused on the formalization and application of the MURM. Specifically, in Chapter 1, the problem is defined and the rationale for applying the MURM is presented with a short history and overview of the methodology. In Chapter 2, the M&S use risk equation is derived from its semantic definition, and the solution to that equation as represented by a three-dimensional surface across the application space is developed. Chapter 3 presents an example of how the MURM is implemented with recommendations and guidance for practitioners. In Chapter 4, a case study based from an actual application is presented that illustrates the efficacy of the methodology to assess a state of risk on a requirement-by-requirement basis while also demonstrating approaches for tailoring V&V activities to reduce risk for the specific intended use of the M&S. Relevant mathematical proofs and details are provided in the appendices.

Finally, MSG-139 held a four-day MURM Workshop in March 2017 during which an application of the MURM was shared with the participants in preparation for the development of more use cases. The goal of the workshop was to prepare the participants to apply the MURM to use cases within their respective domains. Each application of the methodology and the resulting lessons learned serve to increase the confidence that it can be universally applied with success and utility for the user. Expanding the use of the

MURM to other problems and domains will strengthen the methodology itself and is a pathway to more sophisticated implementation strategies. To that end, MSG-139 recommends continuation of use case development by conducting more workshops and establishing a user group through which the participants can share their experiences.

Lignes directrices d’identification, d’analyse et d’atténuation du risque d’utilisation de la modélisation et simulation (M&S)

(STO-TR-MSG-139)

Synthèse

En septembre 2014, le Bureau de soutien à la collaboration scientifique (CSO) de l’Organisation du traité de l’Atlantique Nord (OTAN) a approuvé la formation du Groupe OTAN sur la modélisation et la simulation (NMSG) 139 intitulé « Identification et gestion du risque d’utilisation de la modélisation et simulation (M&S) » [1]. L’objectif principal de ce groupe de travail était de définir et assurer l’application initiale d’une feuille de route afin de soutenir l’élaboration et le déploiement d’une méthodologie, de méthodes et de techniques générales relatives à l’identification et à l’analyse du risque d’utilisation de la M&S et de contrebalancer le risque de M&S par des ressources appliquées à la vérification et la validation (V&V) de la M&S.

Le présent groupe de travail a proposé et formalisé la méthodologie de risque d’utilisation de la M&S (MURM, M&S Use Risk Methodology) [2]. Le groupe a analysé la capacité de la MURM à équilibrer le risque avec le coût et à fournir un moyen d’identification et d’atténuation du risque. Tout en conservant la démarche rigoureuse suivie pendant la mise au point de la MURM, le MSG-139 a commencé par passer en revue la littérature correspondante, à l’aide des informations identifiées et étudiées pendant la revue de littérature de 2009, qui a servi de point de départ [3]. La recherche documentaire actualisée a confirmé que la MURM restait pertinente pour le problème en question et qu’elle était plus mature que toute autre méthodologie. À ce titre, le MSG-139 a choisi de développer la MURM et de l’appliquer.

Dans le contexte de la MURM, le risque d’utilisation de la M&S est défini comme [4] :

> *La probabilité qu’une application inappropriée des résultats de M&S à l’usage particulier prévu ait des conséquences inacceptables pour le décideur.*

La démarche adoptée dans ce travail consiste à traduire cette définition en une logique mathématique servant à calculer le risque d’utilisation de la M&S besoin par besoin. La méthodologie qui en découle s’appuie sur des concepts mathématiques cohérents qui minimisent les biais involontaires. Elle établit une relation explicite avec le processus et les produits de V&V. La MURM peut être utile à plusieurs étapes du processus de développement de la M&S.

Le présent document rapporte les résultats des travaux du MSG-139, y compris la revue de littérature, mais se concentre sur la formalisation et l’application de la MURM. Plus précisément, le chapitre 1 définit le problème et justifie l’application de la MURM avec une brève histoire et une présentation générale de la méthodologie. Le chapitre 2 expose l’équation du risque d’utilisation de la M&S, qui découle de sa définition sémantique, et développe la solution de cette équation, en la représentant par une surface tridimensionnelle dans l’espace applicatif. Le chapitre 3 présente un exemple de mise en œuvre de la MURM, avec des recommandations et des conseils pour les praticiens. Le chapitre 4 expose un cas d’étude tiré d’une application réelle, qui illustre l’efficacité de la méthodologie pour évaluer l’état du risque besoin par besoin, tout en faisant la démonstration de démarches visant à adapter les activités de V&V pour réduire le risque d’utilisation spécifique de la M&S. Les annexes fournissent les preuves mathématiques importantes et les détails.

Enfin, le MSG-139 a organisé un séminaire MURM de quatre jours en mars 2017, pendant lequel une application de la MURM a été présentée aux participants, en vue du développement d'autres cas d'utilisation. Le but du séminaire était de préparer les participants à appliquer la MURM à des cas d'utilisation dans leurs domaines respectifs. Chaque application de la méthodologie et les leçons qui en ont été tirées renforcent la certitude que la MURM peut être universellement appliquée avec succès et de manière utile à l'utilisateur. L'élargissement de l'utilisation de la MURM à d'autres problèmes et domaines raffermira la méthodologie en elle-même et ouvre la voie à des stratégies de mise en œuvre plus sophistiquées. À cette fin, le MSG-139 recommande de poursuivre l'élaboration de cas d'utilisation en réalisant d'autres séminaires et en mettant en place un groupe d'utilisateurs, à travers lequel les participants pourront partager leur expérience.

GUIDELINES FOR MODELLING AND SIMULATION (M&S) USE RISK IDENTIFICATION, ANALYSIS, AND MITIGATION

1.0 ASSESSING MODELLING AND SIMULATION USE RISK

1.1 Background

Models and simulations are developed and employed as enabling technologies to support system analysis, design, test and evaluation, acquisition, training and instruction, and many more areas. Today, a wide variety of Modelling and Simulation (M&S) assets are in use across an even wider range of different application and problem domains. M&S are usually applied when user needs cannot be achieved (e.g., risks, availability) with the actual system or otherwise are achieved more efficiently (e.g., costs, effectiveness) than with the actual system. However, in essence, all M&S assets provide some sort of abstract representation of systems (e.g., entity, phenomenon, process) that are based on different types of approximation. As such, M&S capabilities cannot fully replace the actual system and, more importantly, their usage introduces uncertainties.

Verification and validation (V&V) of M&S are systems and software engineering process areas focused on assessing M&S throughout the life cycle. V&V are implemented to provide the evidence necessary to gain knowledge about M&S assumptions, capabilities, and limitations in relationship to the acceptability criteria. V&V leverage not only systems engineering and software engineering, but also information science, the cognitive and behavioral sciences, and other associated disciplines.

A series of efforts conducted under the umbrella of the North Atlantic Treaty Organization (NATO) Modelling and Simulation Group (MSG), including MSG-054 that resulted in standards and guidance documents for effective V&V of M&S. MSG-054 efforts resulted in approval of the Institute of Electrical and Electronics Engineers (IEEE) Std 1516.4™-2007 Verification, Validation, and Accreditation of a Federation, an Overlay to the High Level Architecture Federation Development and Execution Process in September 2007 and its publication as an international industry standard in December 2007 [5]. In addition to the establishment of the IEEE standard, MSG-054 developed the V&V Composite Model from which to select V&V methods and techniques to match the risk and resource constraints of the V&V efforts while adhering to relevant policies, standards, and guidance [6]. The V&V Composite Model is a superset of the possible activities and the context in which those activities can be tailored into working V&V processes.

MSG-073 achieved standardization of the Generic Methodology for Verification and Validation (GM-VV), illustrated in Figure 1, which provides a generic framework to efficiently develop an argument to justify acceptance and use of identified models, simulations, underlying data, outcomes, and capabilities in the targeted (intended) operational context. The GM-VV has successfully completed the Simulation Interoperability Standards Office (SISO) standardization process through a GM-VV Product Development Group to provide a fully accepted Verification, Validation, and Accreditation (VV&A) guidance document [3]. The purpose of the GM-VV is to provide general applicable guidance for V&V that:

- Facilitates common understanding and communication of V&V within the M&S community;
- Is applicable to any phase of the M&S life cycle (e.g., development, employment, and reuse);
- Is M&S stakeholders' acceptance decision-making process oriented;

- Is driven by the M&S stakeholders' needs and M&S use risks tolerances;
- Is scalable to fit any M&S scope, budget, resources, and use risk thresholds;
- Is applicable to a wide variety of M&S technologies and application domains;
- Will result in traceable, reproducible, and transparent evidence-based acceptance arguments;
- Can be instantiated on enterprise, project, or technical levels alike; and
- Facilitates reuse and interoperability of V&V outcomes, tools, and techniques.

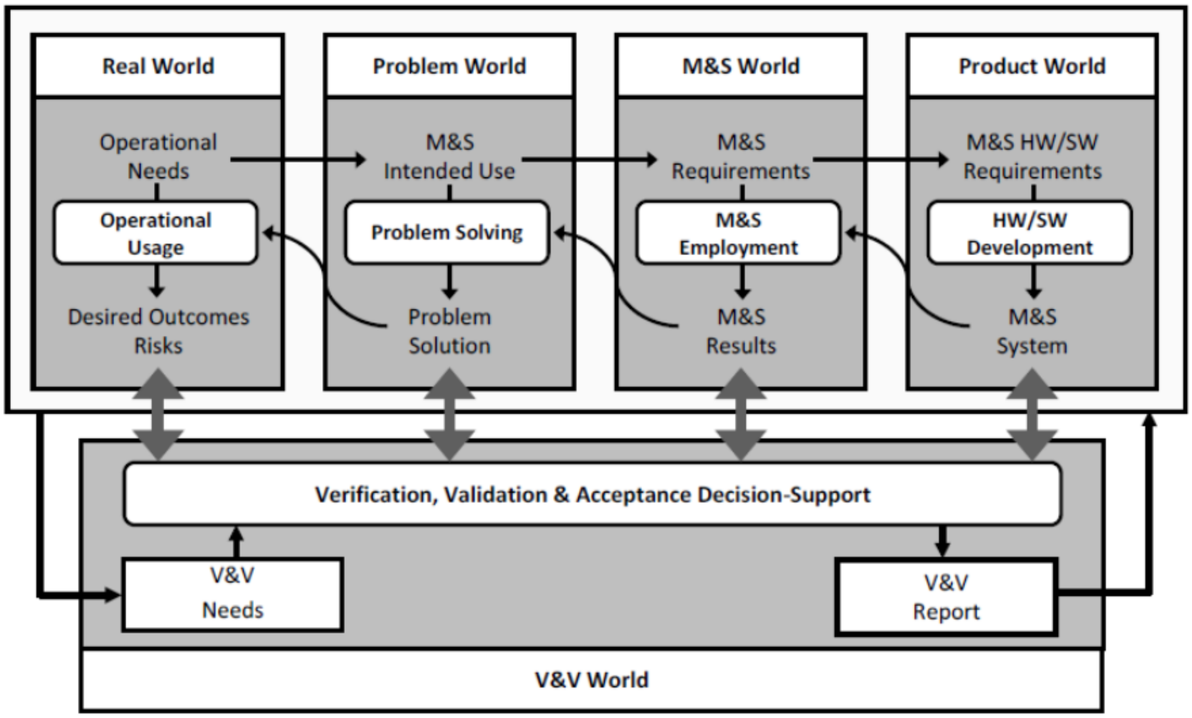

Figure 1: GM-VV Frame of Reference.

In these previous efforts, M&S use risk is recognized and is in fact a driver of the recommendations documented in the guidelines and standards. Despite the consensus in the M&S community on the importance of this topic, there are no accepted methods available for qualification or quantification of M&S use risk that account for project-specific M&S requirements and constraints. In addition, M&S assets and their development processes are increasing in complexity, resulting in a spectrum of risks including M&S use risk. M&S use risks are associated with inappropriate application of M&S results and the consequences of such application for the decision maker.

Risk management relies on assessing the impacts of risks should they be realized, defining methods to mitigate the risk, and evaluating the cost of mitigating the risk. Effective management of risk requires identification of risks and a means by which to balance additional investment to mitigate them. Such an evaluation is made based on an assessment of the likelihood of the realization of the risk and the impact of that realization. When the risks are identified and assessed, mitigation strategies can be developed. A methodology for evaluating M&S use risk can be used to prioritize development objectives, prepare for and respond to changes in resource availability, and tailor V&V activities.

1.2 MSG-139 Objectives, Tasks, and Products

In September 2014, the NATO Collaboration Support Office (CSO) approved the formation of MSG-139, Modelling and Simulation (M&S) Use Risk Identification and Management. The major objective of this task group was to define and deploy a generic methodology with associated methods and techniques for M&S use risk identification and analysis. A common set of complementary and state-of-the-art M&S use risk identification, analysis, and mitigation methodologies facilitate future NATO and national M&S projects with respect to quality, credibility, and utility assurance by providing:

- Generic methods and guidance for M&S Use Risk identification;
- Common understanding and knowledge of M&S use risk issues and solutions;
- A set of methods and techniques for M&S use risk analysis;
- Alternate methods and associated guidelines for tailoring based on M&S use risk rather than cost; and
- M&S use risk identification and analysis solutions that are agnostic with respect to the M&S technology and system life cycle paradigms.

This document reports the results of MSG-139 efforts that satisfy the objectives as stated. Specifically, in Chapter 1, the problem is defined, and the rationale for selecting and applying the M&S Use Risk Methodology (MURM) is presented with a short history and overview of the methodology. In Chapter 2, the M&S use risk equation is derived from its semantic definition, and the solution to that equation, as represented by a three-dimensional surface across the application space, is developed. Relevant mathematical proofs and details are provided in Appendices 1 and 2. Chapter 3 presents an implementation of the MURM with recommendations and guidance for practitioners. In Chapter 4, a use case, based on an actual application, is presented illustrating the efficacy of the methodology to assess a state of risk on a requirement-by-requirement basis while also demonstrating approaches to reduce risk for the Specific Intended Use (SIU) of the M&S.

1.3 Origins of the Modelling and Simulation Use Risk Methodology

In 2009 – 2011, a task to define an approach to provide insightful information regarding the application of M&S to system development was sponsored by the U.S. Deputy Assistant Secretary of Defense for Systems Engineering. The development of the approach was conducted using the scientific method beginning with a thorough literature review of existing approaches that covered 265 documents. The development team then conceived a novel method based on information theory incorporating Bayesian inference drawn from practice that became known as the MURM.

The objective of the MURM is to optimize the use of V&V resources to minimize risk associated with the application of M&S during the development of systems. The MURM leverages existing concepts from numerous communities of interest, as illustrated in Figure 2. Specifically, the MURM utilizes the following concepts:

- The V&V Composite Model and the Validation Process Maturity from the U.S. Department of Defense (DoD) M&S community;
- Confidence ratios, sensitivity analysis, uncertainty quantification, and their communication to decision makers from the design of experiments community;
- Uncertainty quantification, sensitivity analysis, risk assessment, and effective communication to decision makers from NASA-STD-7009;

- Severity categories, probability levels, risk assessment values, and risk acceptable levels from MIL-STD-882;
- Risk-based V&V tailoring, software integrity level, risk matrices, and risk assessment from IEEE-Std-1012-2004;
- Risk-based VV&A, simulation importance, risk assessment, risk matrices, and their effective communication to decision makers, from the U.S. Naval Air Systems Command Battlespace VV&A; and
- Risk-driven software development and the spiral development model from the University of Southern California (USC) Center for Software (SW) Engineering.

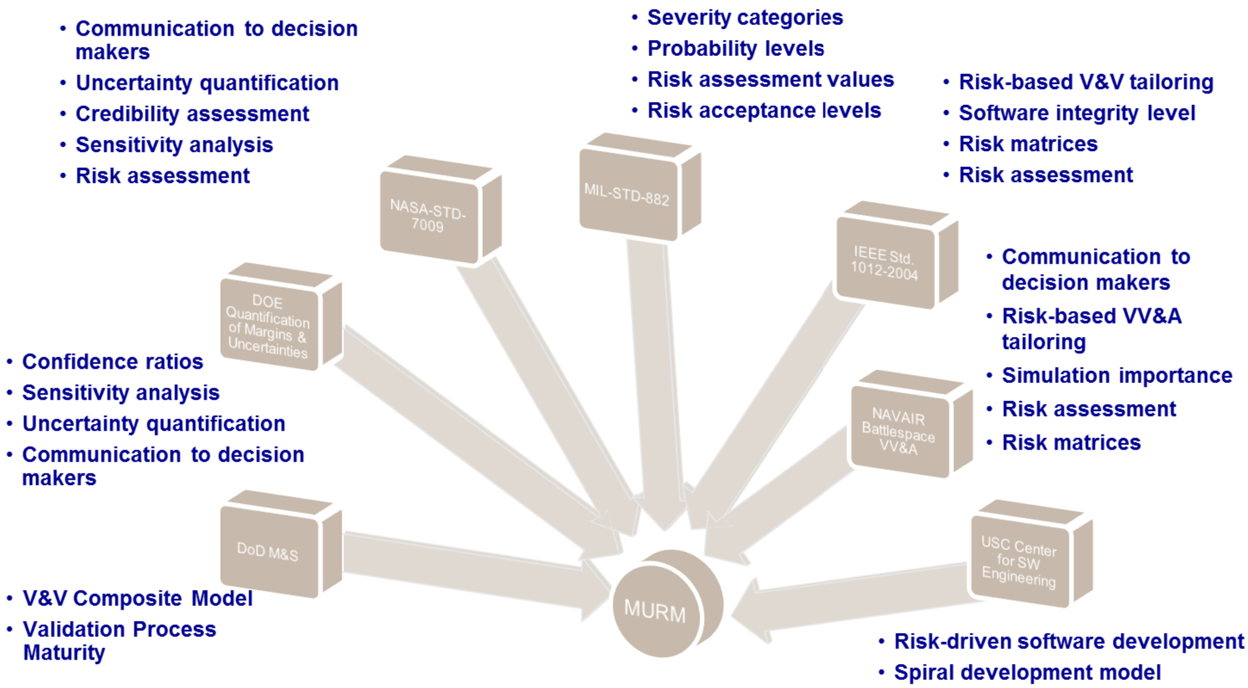

Figure 2: The MURM Builds Upon Existing Concepts.

Over a period of years, the team continued to develop the methodology and test it as use cases became available. In addition, it has been shared with the M&S community of practice and improved by their critical review. Every stage of the development of the methodology was meticulously recorded in the documentation described in Figure 3 [2], [4], [7], [8], [9].

1.4 Updated Literature Survey Summary

In keeping with the rigorous approach utilized in the development of the MURM, MSG-139 initiated its work with a review of the related literature beginning using the information identified and studied during the 2009 literature survey as the starting point. The objectives of the review were to verify that the information from the 2009 survey remains available, relevant, and complete. Initial work was accomplished by the U.S. members at the Johns Hopkins University Applied Physics Laboratory (JHU/APL) and then reviewed and confirmed by the other MSG-139 members. Details of the methods and search results are found in Ref. [10].

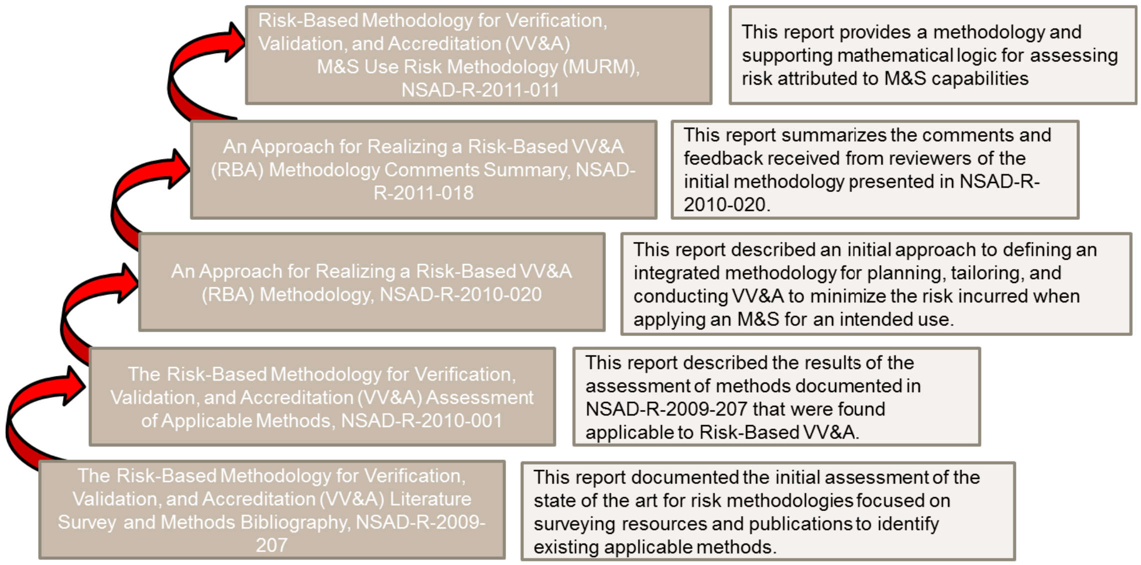

Figure 3: MURM Development Documentation.

The results of these efforts can be summarized as follows:

- There is still concern about risks associated with M&S.
- Information that was available in 2009 remains available, although the location of some information has changed. Information sources should be considered somewhat volatile.
- Some information that was in a hard copy format in 2009 is now available electronically.
- While new sources were discovered that use the 2009 information as references, no significant, new information related to M&S use risk was discovered.

The updated literature search confirmed that the MURM remains relevant to the problem space (i.e., the Problem World) and that it is more mature than any other methodology. Multiple detailed and in-depth discussions within the task group confirmed the viability of the methodology. As such, MSG-139 selected the MURM for further development, implementation, and deployment.

1.5 Workshop

The refinements and implementation of the MURM that are reported in the remaining chapters were reviewed with the MSG-139 participants during a four-day workshop held in March 2017. During the workshop, the theoretical foundations as well as an application of a sample problem were shared with the participants. The goal of the workshop was to prepare the participants to apply the MURM to use cases within their respective domains. Each application of the methodology serves to increase the confidence that it can be used universally with success and utility for the user.

2.0 OVERVIEW OF M&S USE RISK METHODOLOGY

As stated, the objective of the MURM is to optimize the application of V&V resources to minimize M&S use risk. Optimization can be understood colloquially as an act, process, or methodology of making a design, system, or decision as fully perfect, functional, or effective as possible, specifically the mathematical procedures (as finding the maximum of a function) involved in this. Thus, to assert optimization implies quantification of the process, which in this case is quantification of the M&S use risk. Therefore, the first task is to derive a mathematical expression for the M&S use risk.

Semantically, M&S use risk is defined as:

> The probability that inappropriate application of M&S results for the specific intended use will produce unacceptable consequences to the decision maker.

The mathematical approach taken in this work is logically derived from the semantic definition of M&S use risk. The definition can be decomposed into two clauses: (C) inappropriate use of M&S results that leads to (E) unacceptable consequences to the decision maker.

The M&S use risk definition can be represented in mathematical logic as follows:

$$M\&S\ Use\ Risk = P(C \Rightarrow E) \wedge P(C \wedge E) \quad (1)$$

While the solution to the equation is a single quantity (probability), this is not meant to imply that a single M&S use risk value is applied to a simulation as a whole. Rather, an M&S use risk value is computed for each M&S requirement, which is defined at multiple levels. The fundamental trade-off of practicality and precision for applying the methodology is controlled by the granularity the practitioner uses. Applying the methodology against the M&S high level requirements (HLRs) can provide direction that may be especially useful in planning the M&S development prior to the decomposition of the M&S detailed-level requirements (DLRs). Applying the methodology against the DLRs results in a deeper understanding of the source of M&S use risk and actionable plans and recommendations for reducing such risk.

The evaluation of M&S use risk relies on an understanding of the M&S requirements and the V&V process implemented. The MURM surpasses the scope of a standard V&V analysis by mathematically calculating and displaying the M&S use risk for a decision maker on a requirement-by-requirement basis through the assessment of key risk factors.

Figure 4 summarizes a sample application of the MURM that will be detailed in the remainder of this document. In this sample application, M&S use risk is calculated for each M&S requirement based on a series of binary questions regarding the state of each factor. The questions are answered based on documentation from the M&S development and the related V&V activities. The M&S use risk associated with each M&S requirement is plotted on the M&S use risk surface, forming the M&S use risk constellation, which enables the analyst to identify those requirements that are most important to satisfy from the perspective of minimizing M&S use risk.

2.1 Advantages of the MURM

In preparing for use of M&S, decision makers should consider the probabilities that:

1) Capabilities of an M&S are flawed;
2) One or more of the M&S software elements, hardware components, or data are flawed; or
3) The M&S user misunderstands the M&S capabilities.

The MURM provides stakeholders the capability to identify the important aspects of the M&S that when verified and validated will provide the evidence needed to mitigate the risk involved with using M&S results to support decision making. Decision makers can use the MURM across the M&S life cycle to identify and track M&S use risk, make tailoring decisions to focus V&V resource allocations, and make decisions regarding the application of the M&S results.

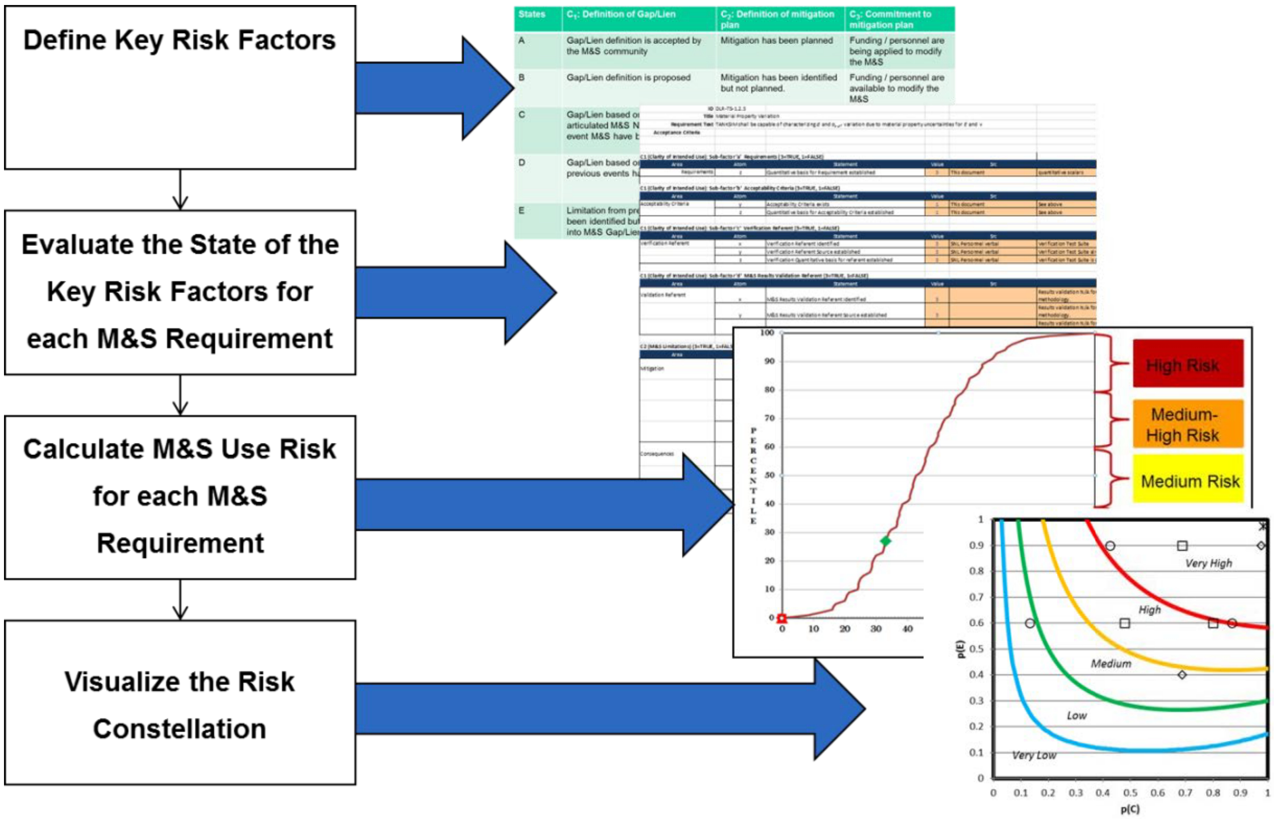

Figure 4: An Implementation of MURM.

The MURM has several characteristics that distinguish it from other processes that consider risk in M&S assessments, as listed here:

1) The MURM is based upon a disciplined and rigorously coherent mathematical process so that its definitions of M&S use risk and other terms have explicit mathematical expressions, and those mathematical expressions are logically consistent and coherent. Methods used have been chosen both for mathematical coherence and to avoid unintended bias (that often creeps into processes) in weightings of factors. The rigorous translation of technical terms into mathematical expressions facilitates automation of many aspects of the M&S assessment process.

2) The MURM is described in general terms because it is a methodology that should be capable of application to any M&S no matter the category, type, domain, or SIU.

3) The MURM is intended to be used by all stakeholders and practitioners associated with either accreditation/acceptance or V&V efforts. It is expected that the MURM will be adopted and used routinely with every M&S, so there will be less risk of misuse of M&S capabilities in decision making.

4) In addition to the mathematical rigor of its processes and concept definitions, the MURM builds upon and reuses other previously published DoD-sponsored VV&A research and credible risk resources.
5) The MURM has multiple purposes that include: (1) to provide a disciplined and mathematically cogent method for assessing risk from M&S use; (2) to facilitate effective and efficient use of V&V resources through the rationale for V&V tailoring; and (3) to provide a mechanism to effectively communicate M&S use risk to the M&S user. All of these impacted the mathematics of the methodology and the purposes are accommodated by the MURM.

These characteristics enable stakeholders to better plan for V&V, monitor the M&S use risk over the development cycle, and adapt activities to respond to increasing M&S use risk.

2.2 Assessing M&S Use Risk to Inform M&S Development

The M&S use risk assessment process begins with a statement of the SIU from which HLRs are derived. The MURM can be applied to HLRs to guide development investment decisions. At this level of granularity, high-level recommendations for investment can be deduced based on the areas where M&S use risk is identified. Such investments may be in the form of model improvements, V&V intensity, or referent development. As the M&S development continues into decomposition of DLRs, the M&S use risk constellation can be revisited to further refine the focus of the M&S development and V&V activities.

2.3 Planning for V&V

V&V contributes to reducing M&S use risk by reducing the uncertainties in the M&S user's knowledge of the M&S capabilities and limitations, thereby reducing the possibility of making use errors. However, complexity and resource limitations may make tailoring of the V&V effort necessary to meet realistic cost and schedule constraints. The MURM supports the development of a V&V Plan tailored to the available resources and schedule as defined by the Accreditation Plan and illustrated in Figure 5 (shaded shapes indicate documentation, and unshaded shapes are activities).

V&V planning begins with determining needed capabilities within the M&S undergoing accreditation/acceptance. The M&S SIU, requirements, acceptability criteria, and associated measures of effectiveness and measures of performance documented in the Acceptance/Accreditation Plan drive the V&V Plan [11]. V&V is planned to provide the evidence needed to address the acceptability criteria. The starting point for defining V&V tasks and activities may be to assume there are no resource constraints. Under such conditions, execution of the plan would produce the most and best possible V&V evidence. An initial evaluation of M&S use risk can be made on the basis of this best-case situation. While the M&S use risk may be low, there will be some level of M&S use risk simply because all M&S exhibit inherent limitations. A second data point in the M&S use risk space can be obtained by constructing a V&V Plan assuming no additional V&V tasks and activities. That is, all evidence will be derived from tasks and activities planned to support the development of the system under test that occur within the timeframe defined by the latest possible due date for the V&V Report as established in the Accreditation Plan. This data point represents the worst-case situation.

Tailoring the V&V effort involves choosing activities and tasks to perform specifically to collect V&V evidence, which techniques to use to execute those activities and tasks, and which parts of the M&S behavior space to explore (i.e., sampling). Having defined the M&S use risk model, the V&V planner can input variations into the model that reflect the impact of activities and tasks, thus producing the series of recommendations for reducing the M&S use risk to an acceptable level. In fact, such information along with the cost associated with additional V&V efforts can be used to execute a trade study of the M&S use risk and cost domains, which are naturally correlated.

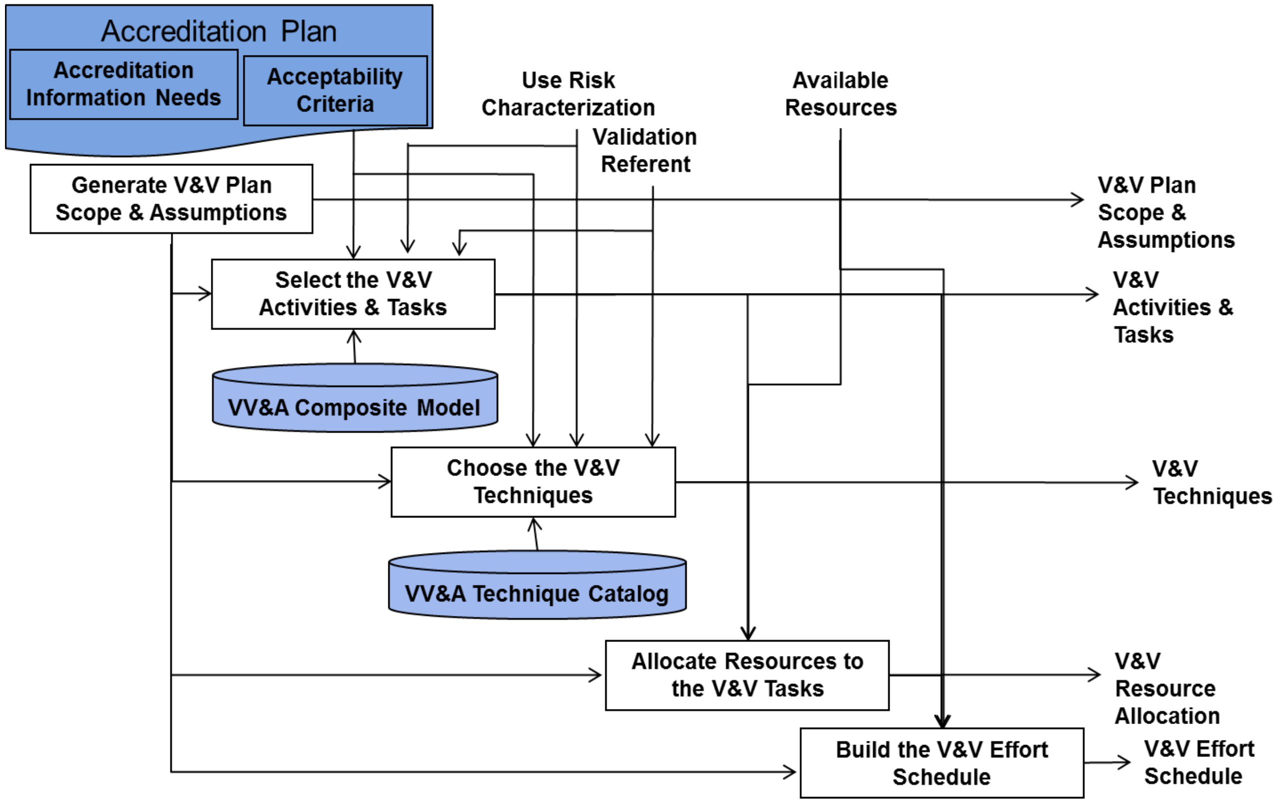

Figure 5: Developing the V&V Plan.

2.4 Using the MURM to Monitor M&S Use Risks

When the V&V Plan is baselined, the M&S use risk evaluation of that plan becomes the M&S use risk baseline. Because the M&S use risk evaluation has been made prior to the development and V&V of the M&S, there are a set of assumptions inherent in the evaluation. As the M&S development and V&V activities are completed, additional quality information becomes available, and the M&S use risk constellation will change. As such, the M&S use risk constellation should be evaluated frequently and certainly at major project milestones.

If the activities and tasks produce the V&V evidence with the quality that was assumed in the evaluation of the baseline M&S use risk, then the final M&S use risk will be within the acceptable M&S use risk defined with the stakeholders. More likely, there will be some failures associated with the activities and tasks, such as cancelled or delayed tests, deferred M&S requirements, or inadequate implementation of configuration management plans. With each such instance, the final M&S use risk becomes less like the desired outcome.

2.5 Using the MURM to Adapt Activities

It is important to remember that change happens in the dynamic execution of programs and plans (e.g., requirements, schedule, and resources could change, etc.). Just as program managers continually manage program risk, M&S use risk should be assessed and managed throughout the M&S life cycle, which affords multiple opportunities to apply the MURM. As program managers make decisions that could affect M&S use risk, V&V practitioners must reassess the impact those program decisions will have on the implementation of

V&V Plans. Any resulting changes in M&S requirements priorities, V&V activities or techniques selected would require a recalculation of M&S use risk. V&V practitioners can then tailor their solutions based on the recalculated M&S use risk and communicate this back to the stakeholders for approval.

Adjustments in priorities and risk will be made as other factors affect the schedule and resources; these are to be expected in any program, experiment, or study. The goal is to mitigate the unexpected when it happens and to provide the M&S user and other stakeholders with a reprioritized requirements listing and the resources required to address the most important requirements at that time.

The MURM enables V&V practitioners to define M&S use risk, identify risk mitigation strategies, and communicate this information to stakeholders effectively. V&V tailoring directly addresses control. V&V practitioners take active steps to minimize M&S use risk by prioritizing M&S requirements and selecting appropriate V&V activities, tasks, and techniques. V&V practitioners must also periodically reassess the potential M&S use risk based on the inevitable changes that will occur during implementation.

3.0 FOUNDATION OF M&S USE RISK METHODOLOGY

This chapter describes the derivation of the M&S use risk equation and the quantification of the M&S use risk through identification and evaluation of the factors that comprise it. Finally, the M&S use risk surface is plotted in a manner to facilitate assessment of the M&S use risk constellation.

The mathematical expressions contained herein generally follow mathematical convention with the exception of probabilities of complex expressions. Probabilities of simple expressions will be shown as lowercase p with a subscript indicating the expression to which the probability applies (e.g., p_1). Probabilities of complex expressions will be shown as uppercase P with the expression to which the probability applies described in parentheses [e.g., $P(A \wedge B)$]. This approach facilitates reading of the document. The notation defined in Table 1 is used throughout this document.

Table 1: Mathematical Expressions.

Symbol	Meaning
$\wedge$	and
$\vee$	or
$\neg$	not
$\Rightarrow$	Implies

3.1 Problem Statement

Recall the semantic definition of the M&S use risk:

> The probability that inappropriate application of M&S results for the specific intended use will produce unacceptable consequences to the decision maker.

Further recall the set of actors and operators inherent in the semantic definition from Section 1.6:

1) The inappropriate application of the M&S.
2) The manifestation of unacceptable consequences.
3) The probability that there is a causal relationship between them.

Given this enumeration as the starting point, the M&S use risk equation is derived below. Let C be the inappropriate application of the M&S. The existence of C is either True or False. The probability that C occurs is expressed as a Bayesian prior probability, p_C. Let E be the unacceptable consequence of the decision. The existence of E is either True or False. The probability that E occurs is expressed as a Bayesian prior probability, p_E. Thus, there are four states that can exist, as shown in Table 2.

Table 2: States for C and E.

Prior Information	State	C	E	P(State)
p_C, p_E	1	T	T	p_1
	2	T	F	p_2
	3	F	T	p_3
	4	F	F	p_4

- State 1 is the condition that the M&S has been used inappropriately, and there are unacceptable consequences. In this state, nothing can be asserted about the causality between C and E.
- State 2 is the condition that the M&S has been used inappropriately. Because no unacceptable consequences occurred, the logical implication is that the inappropriate use did not cause unacceptable consequences.
- In States 3 and 4, the M&S has not been used inappropriately so causality between inappropriate use and unacceptable consequences cannot be ruled out.

Given the prior probability of inappropriate application, p_C, and of the probability of the manifestation of unacceptable consequences, p_E, the M&S use risk definition is predicated in two conditions:

1) The first condition is that C implies E given the prior probabilities, p_C and p_E. This is mathematically written as $(C \Rightarrow E|p_C, p_E)$. Care must be taken to avoid the fallacy of cum hoc ergo proctor hoc,[1] which would assert that coincidence (State 1) is causality.
2) The second condition is that C and E occur at the same time given the prior probabilities. This is mathematically written as $(C \wedge E|p_C, p_E)$. This is State 1 in Table 2, and the probability of this condition is the probability of State 1, p_1.

[1] With it, therefore because of it.

The M&S use risk is the coincidence of the two conditions; that is, both, C implies E, and C and E happen at the same time:

$$M\&S\ Use\ Risk = P(C \Rightarrow E) \wedge P(C \wedge E) \tag{2}$$

Because causality (or relatedness) can only be eliminated in State 2, and the sum of the probability of all the states is one, the total probability of causality is at most the complement of the probability of State 2 (i.e., the sum of the probabilities of States 1, 3, and 4). Combining all of these conditions and applying minimization of the information entropy yield expressions for the probabilities of each state (1 through 4) and, therefore, the probabilities of the conditions. Finally, the M&S use risk (denoted by *UR*) is expressed as the product of the probabilities of the two conditions, which yields the following expression:

$$UR = p_C p_E [1 - p_C + p_C p_E] \tag{3}$$

Formal derivation of this expression is detailed in Appendix A.

3.2 Quantifying the M&S Use Risk

To calculate a value of M&S use risk, values of p_C and p_E must be rationally assigned. Having applied the MURM to more than one M&S domain, the authors assert that the factors influencing the M&S use risk can be summarized and decomposed generically, as shown in Figure 6.

3.2.1 Factor C: Inappropriate Use of the M&S

As defined previously, C is the inappropriate application of the M&S, and p_C is the probability that C occurs. This probability is influenced by three factors: clarity of the intended use, M&S limitations, and confidence in the V&V.

3.2.1.1 Factor C1: Clarity of the Intended Use

The first factor, C1, pertains to the clarity of the M&S SIU. Lack of clarity in specifying the SIU increases the probability of inappropriate application of the M&S. The SIU is typically defined in a planning document, such as an Accreditation Plan or V&V Plan, and should be accompanied by all of the necessary ancillary SIU clarification products, such as HLRs, DLRs, acceptability criteria, and identified referents for requirements verification and M&S results validation.

3.2.1.2 Factor C2: M&S Limitations

The second factor, C2, pertains to known limitations in the M&S with respect to specific DLRs, their consequences relative to the SIU, and their current mitigation status. Limitations, as defined for the purpose of the MURM, are restrictions in the ability of the M&S to represent the simuland with sufficient fidelity over the SIU problem space. M&S limitations are typically documented in V&V Reports. If they are not fully mitigated, they increase the probability of misapplication of the M&S results to varying degrees, depending on the severity of the limitation consequences to the SIU.

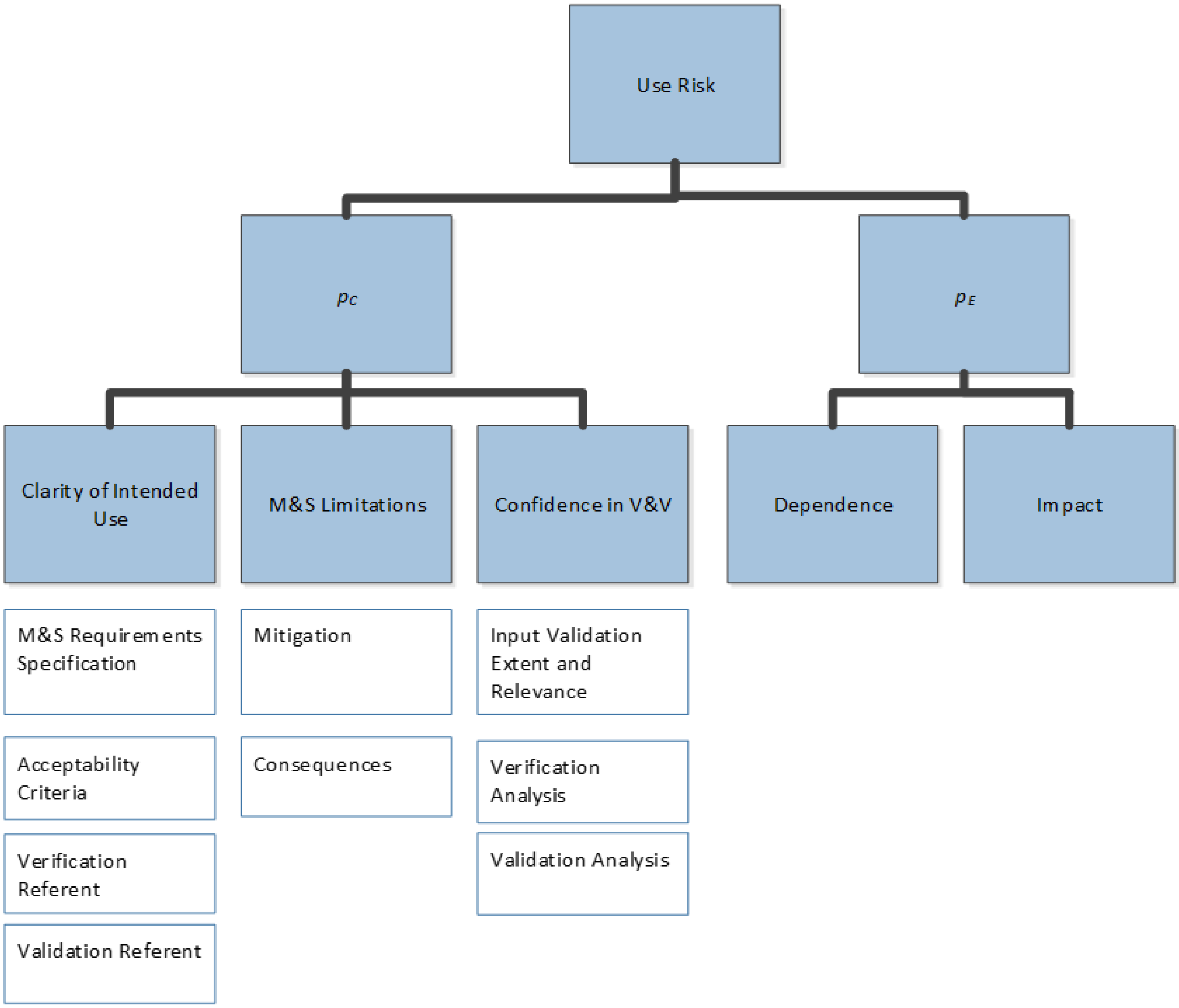

Figure 6: MURM Influence Factors.

3.2.1.3 Factor C3: Confidence

The third factor, C3, addresses the level of confidence in the V&V analysis itself, as opposed to the V&V analysis results. Deficiencies in the scope and rigor of the V&V analysis relative to a specific DLR, or in the applicability of the V&V testing referents and testing frameworks, increase the probability of misapplication of M&S results. In other words, even if the analysis of the V&V test results indicates M&S applicability to the SIU, there is a greater probability of M&S results misapplication and possible unacceptable consequences to the decision maker if there are reasons to have reduced confidence in the V&V analysis itself.

V&V analysis is comprised of three distinct activities that are the sub-factors for C3: input validation, requirements verification, and results validation. Input validation analysis involves determining whether the M&S inputs associated with a DLR are complete and relevant for the SIU, from a valid and authoritative source, and correctly implemented in the M&S. Requirement verification analysis involves determining whether all facets of the DLR were implemented fully and correctly in the M&S. Finally, M&S results validation analysis involves determining the fidelity of M&S results pertinent to the DLR relative to the real-world system and assessing whether that fidelity is sufficient for the SIU.

3.2.2 Factor E: Effect of Unacceptable Consequences

As defined, E is the manifestation of unacceptable consequences of decisions based on the M&S results, and p_E is the probability that E occurs. This probability is influenced by two sub-factors: dependence and impact.

Dependence refers to the degree to which the decision maker is basing the decision on the M&S results. Rarely would a decision maker place full confidence in a simulation result to make critical decisions. Usually, a decision maker is presented with a body of knowledge gathered through test and evaluation, and simulation upon which to rely to make the decision. If the situation developed that the M&S result was the only available information upon which to base the decision, then the influence of this factor would increase.

Impact refers to the degree to which the requirement being addressed by the requirement being evaluated represents a critical aspect of the M&S. For example, M&S requirements that address the ease with which the operator can start or stop a simulation will have less influence on the credibility of the M&S results than a requirement to faithfully represent a function of the system being modelled. If the SIU is performance assessment, failing to satisfy either requirement could adversely impact the ability of the analyst to produce M&S results in a timely manner; however, the latter would have direct impact on the credibility of the M&S results, and the influence of this factor would increase.

The factors presented earlier have been refined and applied to several use cases and found to be generally applicable across domains. Additional use cases would confirm the general applicability of these factors and their related states.

3.3 Calculating and Visualizing M&S Use Risk

Each factor or sub-factor is decomposed into a series of statements referred to as atoms. The truth of each of the statements and the set of possible states of the sub-factors define the influence of each factor as interpreted in the field of information theory. The technique for evaluating the influence of each factor is presented in Appendix B and has been applied to the taxonomy of factors presented here to result in a consistent method for evaluating the probabilities associated with each factor. The numeric results of these evaluations are shown in the tables in Appendix B.

The value of p_C is calculated from the values of the probabilities of the independent sub-factors and applying the sum rule for probability, as follows:

$$p_C = p_{C1} \vee p_{C2} \vee p_{C3} = p_{C1} + p_{C2} + p_{C3} - (p_{C1}p_{C2} + p_{C1}p_{C3} + p_{C2}p_{C3}) + p_{C1}p_{C2}p_{C3} \quad (4)$$

Having quantified p_C and p_E, the M&S use risk can be evaluated directly from the previously derived equation:

$$UR = p_C p_E [1 - p_C + p_C p_E] \quad (5)$$

To better understand the nature of this equation, the M&S use risk can be plotted in a three-dimensional space over the expected range of p_C and p_E, yielding the smooth surface shown in Figure 7. The M&S use risk associated with each requirement can be calculated and plotted onto the surface. To facilitate assessment of the

M&S use risk constellation, lines of equal M&S use risk (isoclines) can be plotted onto the surface, as shown in Figure 8. To further facilitate visualization and comparison, the surface, isoclines, and individual evaluations can be projected onto the $p_C - p_E$ plane, yielding the view shown in Figure 9. Thus, evaluation of M&S use risk for a particular requirement can be assessed relative to the M&S use risk for all other requirements. The behavior of Figure 9 is visually counterintuitive in regions of high values of p_C and very low values of p_E (e.g., near the marker labelled "2"). The explanation of this phenomenon can be drawn directly from the definition of entropy in information theory and is presented in Appendix C.

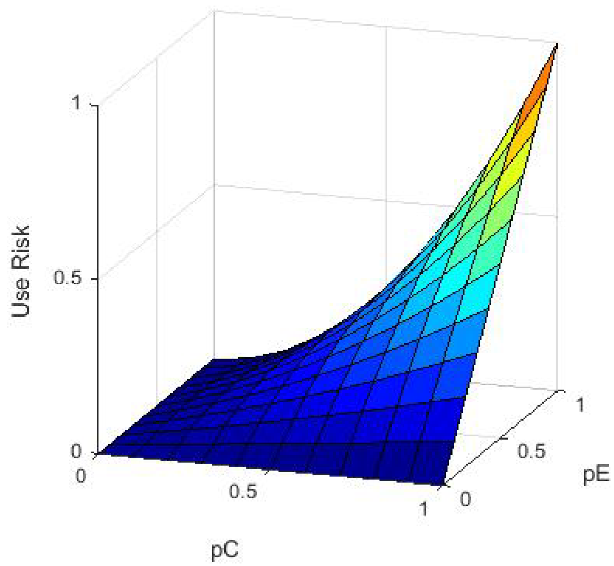

Figure 7: M&S Use Risk Surface.

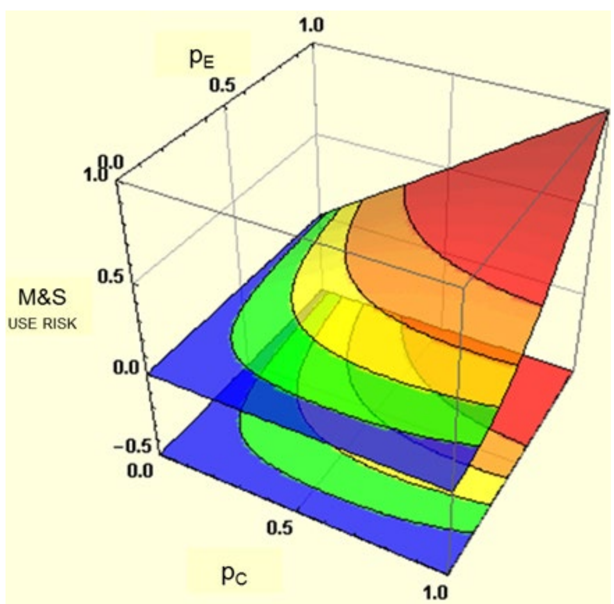

Figure 8: M&S Use Risk Surface with Isoclines.

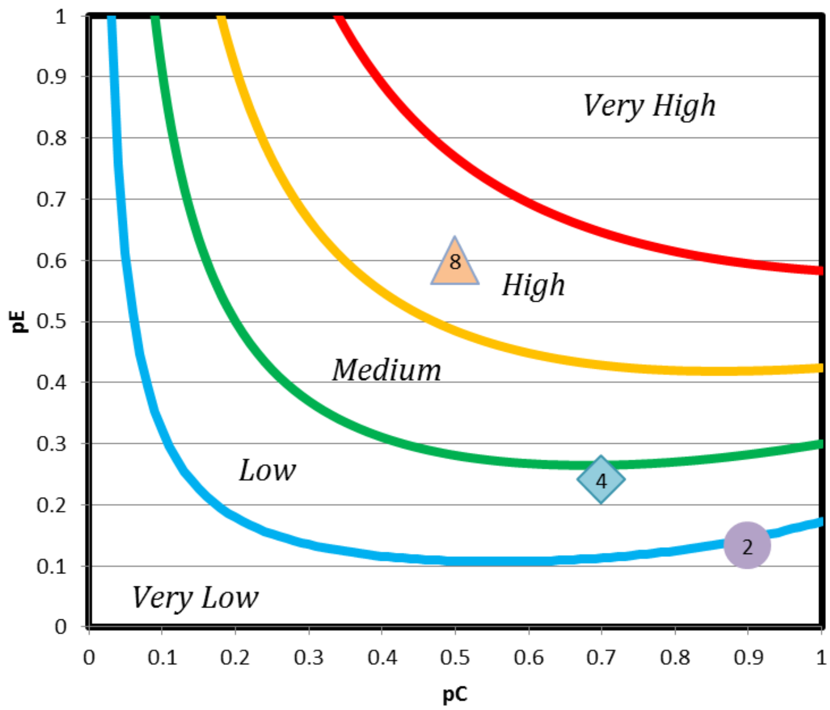

Figure 9: Two-Dimensional Projection of the Risk Constellation.

4.0 SAMPLE APPLICATION

4.1 Link to the M&S Development and V&V Activities

Basic systems or software engineering principles are applied to M&S development or modification and conform to the system engineering model of development depicted in Figure 10. Requirements to build or modify M&S are identified; transformed into a conceptual model, various specifications, and design drawings; and then implemented into software and hardware components. V&V, shown in Figure 10 as the elements of the GM-VV developed by NMSG Task Group 073, are the engineering practices implemented and documented to ensure requirements are traceable throughout intermediary development products and alignment with the system engineering activities, according to the development paradigm selected for the project [3]. A third layer representing risk management throughout the development is linked to the V&V activities through the timely application of the MURM. The end result of the M&S development or modification processes is acceptance-testing leading to acceptance of the delivered product by the stakeholder.

No matter the type of M&S to be used or the reason that it requires V&V to be performed, the V&V processes must begin with a statement of the M&S SIU and the specification of the M&S requirements. Providing this information is the responsibility of the stakeholders – the people with the need to use M&S to support their program, mission, objectives, analysis, experiment, training, etc. The requirements for the M&S are derived from the SIU. In the case of use of an existing M&S product, the needs of the stakeholders are compared against the requirements to which the M&S was originally developed along with the information in the available development products. The output of those comparisons defines the requirements for a modification to the existing product.

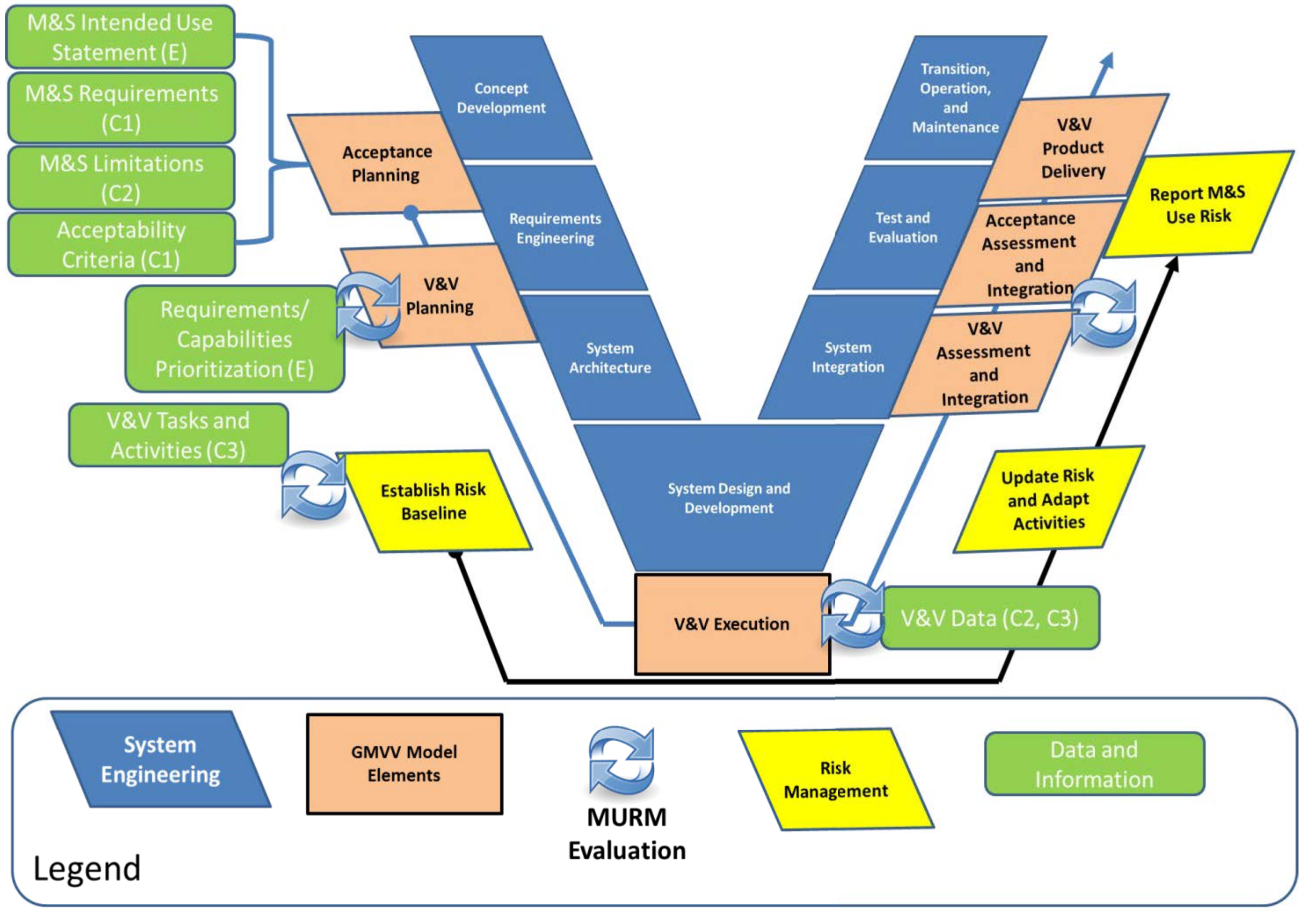

Figure 10: M&S Use Risk Assessment Throughout the M&S Development Life Cycle.

The M&S SIU and M&S requirements together represent the information needed by the accreditation authority to initiate accreditation planning and produce the acceptability criteria documented in the Accreditation Plan. In planning the V&V, the stakeholders must communicate their priorities for use of the M&S. The result should be a prioritized listing of the requirements of the M&S that the V&V agent uses to tailor the V&V Plan. Communication of the priorities is important because the M&S user and other stakeholders must make resourcing decisions based upon the priorities and available V&V resources. Once decisions are made, then V&V planning can continue, estimates of the required resources and schedule for the highest priorities can be determined, and M&S use risk can be evaluated. This first evaluation of the M&S use risk establishes the M&S use risk baseline.

The V&V Report documents the results of executing the V&V Plan, captures the changes made during execution, and produces a map of the M&S capabilities and limitations to enable M&S users to shape the use of the M&S to take advantage of its capabilities and avoid the pitfalls of its limitations. The V&V Report communicates the residual M&S use risk associated with each M&S requirement explored during execution of the V&V Plan.

4.2 Documentation as Input

Figure 10 identified several pieces of data and information required to complete a MURM evaluation. Assuming the M&S development is adhering to systems engineering best practices, these data and information should be available in the program/project, M&S development, and V&V documentation.

The M&S SIU statement may be recorded in a letter of instruction from the M&S user, M&S proponent, or a Project Management Plan. The SIU statement should also be documented in the Accreditation Plan and/or the V&V Plan.

The M&S requirements can be found in requirements documents and may be maintained in a requirements management tool. The M&S Development Plan (or similar document) may also include a requirements burn-down plan. M&S requirements should also be documented in the Accreditation Plan and/or the V&V Plan and should form the basis of the V&V and Integration Test Plans. A useful resource is a requirements traceability matrix, such as in Figure 11. If the M&S is being developed using an evolutionary methodology (such as Agile), the M&S requirements will not be fully developed at the outset of the development project. The MURM should be revisited as the M&S requirements evolve.

Example Requirements Traceability Matrix

Test Results summarized from documented test Report

Req #	Requirement Derived From	Requirement Description	V&V Layer	Source of V&V Test Result(s)	V&V Test Description	Test Procedure(s)	Test Result
001	Reference[2] (120), Reference[3] (Table 3-2), Reference[4] (Section 1.4)	The Simulator shall multicast surveillance data for up to 30 static aircraft in the format specified.	3	Reference[9]	**1-01**: Verify the Simulator outputs surveillance data for up to 30 aircraft	Preconditions:Ensure the Simulator, Sensor System, and the measurement tool on the Simulator computer are running.	✔
						1-01.1 On the Simulator GUI display/Ship System Outputs tap, verify the check-box of Enable & the check-box of Ship Position Input is checked under the Aircraft Scenario.	✔
						1-01.2 On the Simulator GUI display, verify the 30 A/C with Surv Output is selected under the Scenario Option.	✔
						1-01.3 On the measurement tool, observe the 30 surveillance aircrafts data being outputted by the Simulator at the data rate of 4 Hz.	✔
						1-01.4 On the measurement tool, observe the 30 surveillance aircrafts positioning (Latitude, Longitude, and altitude), with the appropriate identifier codes as specified by the 30 static surveillance aircrafts scenario parameters.	✔

Figure 11: Exemplary Requirements Traceability Matrix.

M&S limitations are described in U.S. MIL-STD-3022 as:

> "the known constraints and limitations associated with the development, testing, and/or use of the M&S. These constraints and limitations may be introduced as a result of an ongoing development process or may result from information garnered in previous V&V efforts. Limiting factors include constraints on

M&S capability as well as constraints associated with M&S testing that may result in inadequate information (e.g., inadequate resources, inadequate technical knowledge and subject matter expertise, unavailable data, inadequately defined M&S requirements and methodologies, and inadequate test environments) to support the M&S assessment process [11]."

All M&S have limitations because they are representations of real-world entities and not the entities themselves. Some M&S limitations are conceded at the beginning of the development process as aspects of the real-world entity that will not be represented. These should be recorded in the Accreditation Plan and the V&V Plan. Others arise during development due to resource constraints or technology limitations. Some limitations and constraints will be induced by the inherent assumptions in the model and the level of detail incorporated and may only be discovered or quantified during validation analysis. Such limitations should be recorded in the V&V Report.

Acceptability criteria are set during accreditation planning and reflect the M&S user's need for accuracy and utility of the M&S. Methods for defining and deriving acceptability criteria vary widely as does the party responsible for deriving the criteria. This role may be undertaken by the accreditation agent, the accreditation authority, a user, a stakeholder, members of the V&V team, or a combination of any or all these parties. As noted by Harmon and Youngblood, the acceptability criteria "need to be acceptable to all of the players in a simulation development, modification or application" [12]. This can be facilitated by involving all application domain SMEs and the stakeholders in deriving the candidate acceptability criteria. The accreditation authority has the final responsibility for setting the criteria, which are then recorded in the Accreditation Plan.

Tailoring is a primary task during V&V planning. The first level of tailoring occurs when determining priorities of the M&S requirements within the scope of the overall V&V effort as defined by the SIU. Focusing on the M&S requirements that are most critical to the SIU through prioritization will provide the richest set of evidence to support either an acceptance or an accreditation decision. There are multiple ways to prioritize M&S requirements (e.g., binning, structured relationships), and the prioritization should be contained within the project documentation. Regardless of the method by which the prioritization is set, the activity should be supported by the M&S user, the developers, and related SMEs.

Documenting the V&V tasks and activities is the purpose of the V&V Plan document, and it should be used to assess the related MURM factors. The enumeration of tasks and activities will include the identification of the referent as the "Basis of Comparison" in Appendix C of the V&V Plan [11]. Similarly, the V&V Report is intended to summarize the V&V evidence, which should be archived and made available to the V&V agent and will also be used to support MURM evaluations.

For all MURM evaluations, rationale and references to data and information used for the scoring of each factor for each requirement should be recorded to assure consistency and to facilitate follow-on evaluations. The state of these data and information will change over the course of the development of the M&S. Each state change will impact the M&S use risk constellation and re-evaluation should be undertaken appropriately.

4.3 Workbook Implementation

The MURM was prototyped as a Microsoft Excel workbook. Each workbook is comprised of a series of True/False statements that were derived from the factors and sub-factors identified in Chapter 3. Table 3, Table 5 and Table 7 provide the workbooks for factors C1, C2, and C3, respectively. The shaded (orange) cells in each table indicate input required from the V&V agent. The V&V agent provides the answers to the questions for each requirement based on the M&S and V&V documentation and consultation with SMEs. Atom statement evaluations are generally True/False evaluations made according to the guidance in the tables following each workbook image. Table 4, Table 6, and Table 9 provide guidance for completing the C1, C2, and C3 workbooks, detailing specific considerations one should account for when assessing each MURM atom for each requirement. Each workbook table also includes space for the V&V agent to record the Source (Src) of documentation used to make each evaluation and for comments to facilitate consistent re-evaluation of the M&S use risk throughout the M&S life cycle. The workbook provides the underlying mathematics to calculate the M&S use risk associated with each requirement and a template for producing a plot of the M&S use risk constellation. Some atoms are mutually exclusive, and such situations are handled in the workbook logic.

4.3.1 Factor C: Inappropriate Use of the M&S

4.3.1.1 C1: Clarity of Intended Use

Table 3: Workbook Table for Factor C1.

C1 (Clarity of Intended Use): Sub-factor 'a' Requirements (3=TRUE, 1=FALSE)

Area	Atom	Statement	Value	Src	Comment
Requirements	z	Quantitative basis for Requirement established			

C1 (Clarity of Intended Use): Sub-factor 'b' Acceptability Criteria (3=TRUE, 1=FALSE)

Area	Atom	Statement	Value	Src	Comment
Acceptability Criteria	y	Acceptability Criteria exists			
	z	Quantitative basis for Acceptability Criteria established			

C1 (Clarity of Intended Use): Sub-factor 'c' Verification Referent (3=TRUE, 1=FALSE)

Area	Atom	Statement	Value	Src	Comment
Verification Referent	x	Verification Referent Identified			
	y	Verification Referent Source established			
	z	Verification Quantitative basis for referent established			

C1 (Clarity of Intended Use): Sub-factor 'd' M&S Results Validation Referent (3=TRUE, 1=FALSE)

Area	Atom	Statement	Value	Src	Comment
Validation Referent	x	M&S Results Validation Referent Identified			
	y	M&S Results Validation Referent Source established			
	z	Validation Quantitative basis for referent established			

Table 4: Guidance for Evaluation of Factor C1.

Sub-Factor	Atom Statement (T/F)	Considerations
Requirements	Quantitative basis for requirement established	Does the description of the requirement contain any quantitative or mathematical descriptors (e.g., counts, percentages, functional equations) to define the requirement or is it comprised only of qualitative descriptors? Requirements with a quantitative basis decrease M&S use risk by defining the SIU of the M&S with more precision than purely qualitative requirements.
Acceptability Criteria	Acceptability criteria exists	Does the requirement have an associated acceptability criterion? An acceptability criterion is a standard that the M&S and its associated data must meet, with respect to the requirement, such that it can be accredited for the SIU. Lack of an acceptability criterion increases M&S use risk because the requirement has no definitive criteria to determine whether it has been met. Without acceptability criteria, the required fidelity of the M&S for the SIU is essentially undefined.
	Quantitative basis for acceptability criteria established	Does the description of the acceptability criterion contain any quantitative mathematical or statistical limits? Acceptability criteria with a quantitative basis reduce M&S use risk by providing precise criterion to determine whether the M&S performance, relative to the associated requirement, is sufficient for the SIU. Commonly found descriptions such as "The M&S is acceptable if evidence shows < repeat of qualitative requirement verbiage>" are not informative and not useful for reducing M&S use risk.
Verification Referent	Verification referent identified	Is there a documented referent for the verifying satisfaction of the M&S requirement? Referents are typically identified in the M&S V&V Plan and include relevant portions of build, capability, or interface specifications for the M&S or tactical system being simulated. For complex requirements that span a lot of solution space, referents may also be other models independently written to the same specifications as the M&S. Clearly identified verification referents decrease M&S use risk because they provide more authoritative, detailed, and precise referent information than just the requirement description itself.
	Verification referent source established	Has the identified referent actually been instantiated? In other words, has the specification document or independent model to be used as the basis for verification comparison actually been written or created? SIU clarity risk is not mitigated if the identified referent does not actually exist, even if completion in the required timeframe for verification analysis is highly likely. This is because the adequacy of the referent for use in verification analysis cannot be assured until it is actually instantiated.

Sub-Factor	Atom Statement (T/F)	Considerations
	Verification quantitative basis for referent established	Does the description of the referent or planned verification methodology contain any quantitative or mathematical bounds defining how the M&S output will be compared to the referent for the purpose of requirement verification or is it comprised only of qualitative descriptors? Requirements with a quantitative basis defined for the verification analysis decrease M&S use risk by defining the scope of the requirement relative to the SIU of the M&S with more precision than those with a purely qualitative basis.
Validation Referent	M&S results validation referent identified	Is there a documented referent for validating M&S results, pertinent to the requirement, as being representative of real-world behavior? Validation referents are typically identified in the M&S V&V Plan. They may include such items as SME opinion, benchmark simulation results, real-world calibration test results, or ideally, real-world test results from representative operational scenarios. Lack of clearly identified results validation referents increases M&S use risk because no validation against real-world behavior can be performed. Thus, although the requirement may be fully verified, it is impossible to know if the requirement itself, as implemented, is valid for the SIU.
	M&S results validation referent source established	Has the identified validation referent been instantiated? If the referent is to be SME opinion, has a SME been identified and tasked with reviewing the validation evidence? If the referent is to be a benchmark simulation, has it been acquired, and have the referent scenario results been generated? If the referent is to be calibration or live test results, has the test been completed and the results compiled such that they can be used for comparison against M&S results? SIU clarity risk is not mitigated if the identified referent does not exist, even if completion in the required timeframe for validation analysis is highly likely. This is because the adequacy of the referent for use in validation analysis cannot be assured until it is instantiated.
	Validation quantitative basis for referent established	Does the description of the referent or planned validation methodology for the requirement contain any quantitative or mathematical bounds defining how the M&S output will be compared to the referent for the purpose of results validation, or is it comprised only of qualitative descriptors? Requirements with a quantitative validation basis decrease M&S use risk by defining the fidelity of the M&S results necessary for the SIU with more precision than those with a purely qualitative basis.

4.3.1.2 C2: M&S Limitations

Table 5: Workbook Table for Factor C2.

C2 (M&S Limitations) (3=TRUE, 1=FALSE)					
Area	Atom	Statement	Value	Src	Comment
Mitigation	s	Current mitigation status of all known limitations pertinent to the requirement is 'not required'			
	t	Current mitigation status of all known limitations pertinent to the requirement is 'complete'			
	u	Current mitigation status of all known limitations pertinent to the requirement is 'partially complete'			
	v	Current mitigation status of all known limitations pertinent to the requirement is 'impossible'			
Consequences	w	Consequences to the intended use of all known M&S limitations pertinent to the requirement are neglible			
	x	Consequences to the intended use of all known M&S limitations pertinent to the requirement are minor			
	y	Consequences to the intended use of all known M&S limitations pertinent to the requirement are serious			
	z	Consequences to the intended use of all known M&S limitations pertinent to the requirement are grave			

Table 6: Guidance for Evaluation of Factor C2.

Sub-Factor	Atom Statement (T/F)	Considerations
Mitigation	Not required	Answer is "True" if there are no known limitations or if the consequences to a known limitation are no worse than "negligible." Otherwise, answer is "False."
	Complete	Answer is "True" if known limitations have been mitigated through code fixes, bias adjustments, etc., and there is documented proof of the effectiveness of the fixes. Otherwise, answer is "False."
	Partially complete	Answer is "True" if known limitations have been partially mitigated or claims of full mitigation have been made, but there is no documented proof of effectiveness. Otherwise, answer is "False."
	Impossible	Answer is "True" if there is no known mitigation strategy or if a known strategy will definitely not be implemented due to cost, schedule, or priority constraints. Otherwise, answer is "False."
Consequences	Negligible	Answer is "True" if there are no known limitations or if the consequences from a limitation, in the context of the overall problem space covered by the M&S SIU, are negligible. Otherwise, answer is "False."
	Minor	Answer is "True" if the overall consequences to the SIU from the limitation are minor. This could be a limitation that encompasses the entire problem space with a minimal impact or a limitation with a more serious impact but only within a small portion of the problem space. Answer is "False" if none of the aforementioned conditions are satisfied.
	Serious	Answer is "True" if the overall consequences to the SIU from the limitation are significant. Otherwise, answer is "False."
	Grave	Answer is "True" if the overall consequences to the SIU from the limitation are significant and will affect most of the problem space. Otherwise, answer is "False."

4.3.1.3 C3: V&V Confidence

Table 7: Workbook for Factor C3.

C3 (Confidence) Sub-factor 'a': Input Validation Analysis (3=TRUE, 1=FALSE); User input "None", "Partial", or "Full" for Extent and Relevance

Area			Atom	Statement	Value	Src	Comment
Scope	Extent		u	Scope Meets Objective			
	Relevance		v	Scope Meets Threshold			
Referent			w	Referent Meets Objective			
			x	Referent Meets Threshold			
Methodology			y	Methodology Meets Objective			
			z	Methodology Meets Threshold			

C3 (Confidence) Sub-factor 'b': Verification Analysis (3=TRUE, 1=FALSE); User input "None", "Partial", or "Full" for Extent and Relevance

Area			Atom	Statement	Value	Src	Comment
Scope	Extent		u	Scope Meets Objective			
	Relevance		v	Scope Meets Threshold			
Referent			w	Referent Meets Objective			
			x	Referent Meets Threshold			
Methodology			y	Methodology Meets Objective			
			z	Methodology Meets Threshold			

C3 (Confidence) Sub-factor 'c': M&S Results Validation Analysis (3=TRUE, 1=FALSE); User input "None", "Partial", or "Full" for Extent and Relevance

Area			Atom	Statement	Value	Src	Comment
Scope	Extent		u	Scope Meets Objective			
	Relevance		v	Scope Meets Threshold			
Referent			w	Referent Meets Objective			
			x	Referent Meets Threshold			
Methodology			y	Methodology Meets Objective			
			z	Methodology Meets Threshold			

4.3.1.3.1 Scope Assessment

The assessments of scope for input validation, requirements verification, and M&S results validation are assessments of both the ***extent*** of the analysis and the ***relevance*** of the analysis to the M&S SIU.

Assessing extent, in the context of input validation analysis, requires a consideration of how many of the inputs relevant to the requirement in question have been analyzed for validity. Assessing the extent of the requirements verification involves consideration of how many of the aspects, or facets, of the requirement in question were verified. Assessing the extent of the M&S results validation involves consideration of how many of the M&S outputs, pertinent to the requirement, were assessed for validity against a real-world referent. The extent assessment should be characterized by one on the three following words: “full,” “partial,” or “none.”

Unlike extent, assessment of the relevance of an analysis does not vary by the type of analysis. For all three types of analysis, assessing the relevance involves consideration of whether the evidence is actually relevant to the M&S SIU. In other words, was the verification and validation evidence produced with a version of the M&S comparable in both function and performance to the version of the M&S that will be used for the SIU? Items to consider when assessing relevance include the age of the analysis evidence, the deltas between the analyzed and SIU software version numbers, and version description documentation. Version description documents, when available, can be used to confirm comparable function and performance, relative to the requirement in question, even when analysis evidence is rather old, or the software version number deltas are large. Similar to extent, the relevance assessment should ultimately be characterized by one of the three following words: “full,” “partial,” or “none.”

Table 8 provides the rules for combining the assessments for extent and relevance into a single scope assessment. This process has been automated in the workbook implementation.

Table 8: Extent/Relevance Combination Rules for Scope Assessment.

Extent	Relevance	Scope
None	–	Scope Meets Objective = False Scope Meets Threshold = False (i.e., does NOT meet threshold)
Partial	None	
Full	None	
Partial	Partial	
Partial	Full	Scope Meets Objective = False Scope Meets Threshold = True (i.e., meets threshold)
Full	Partial	
Full	Full	Scope Meets Objective = True Scope Meets Threshold = True (i.e., meets objective)

Table 9: Guidance for Evaluation of Factor C3.

Sub-Factor	Atom Statement (T/F)	Considerations
Input Validation Analysis	Scope Meets Objective	Was the extent and relevance of the input validation analysis fully sufficient (i.e., meets objective), barely sufficient (i.e., meets threshold), or insufficient (i.e., does NOT meet threshold)? See Table 8.
	Scope Meets Threshold	
	Referent Meets Objective	Is the input referent, i.e., source of the inputs, fully sufficient for the M&S SIU? If so, answer is "True;" otherwise, answer is "False."
	Referent Meets Threshold	Is the input referent, i.e., source of the inputs, barely sufficient (answer "True") or insufficient (answer "False") for the M&S SIU?
	Methodology Meets Objective	Was the input validation methodology used fully sufficient for the M&S SIU? If so, answer is "True;" otherwise, answer is "False."
	Methodology Meets Threshold	Was the input validation methodology barely sufficient (answer "True") or insufficient (answer "False") for the M&S SIU?
Verification Analysis	Scope Meets Objective	Was the extent and relevance of the verification analysis fully sufficient (i.e., meets objective), barely sufficient (i.e., meets threshold), or insufficient (i.e., does NOT meet threshold)? See Table 8.
	Scope Meets Threshold	
	Framework Meets Objective	Was the testing framework used for the verification analysis fully relevant to the framework that will be used for the M&S SIU? If so, answer is "True;" otherwise, answer is "False."
	Framework Meets Threshold	Was the testing framework used for the verification analysis somewhat relevant to the framework to be used for the M&S SIU (answer "True") or irrelevant (answer "False")?
	Methodology Meets Objective	Was the verification methodology used fully sufficient for the M&S SIU? If so, answer is "True;" otherwise, answer is "False."
	Methodology Meets Threshold	Was the verification methodology barely sufficient (answer "True") or insufficient (answer "False") for the M&S SIU?

Sub-Factor	Atom Statement (T/F)	Considerations
Output Validation Analysis	Scope Meets Objective	Was the extent and relevance of the results validation analysis fully sufficient (i.e., meets objective), barely sufficient (i.e., meets threshold), or insufficient (i.e., does NOT meet threshold)? See Table 8.
	Scope Meets Threshold	
	Referent Meets Objective	Is the results referent, i.e., source of the real-world data for results comparison, fully sufficient for the M&S SIU? If so, answer is “True;” otherwise, answer is “False.”
	Referent Meets Threshold	Is the results referent, i.e., source of the real-world data for results comparison, barely sufficient (answer “True”) or insufficient (answer “False”) for the M&S SIU?
	Methodology Meets Objective	Was the results validation methodology used fully sufficient for the M&S SIU? If so, answer is “True;” otherwise, answer is “False.”
	Methodology Meets Threshold	Was the results validation methodology used fully sufficient for the M&S SIU? If so, answer is “True;” otherwise, answer is “False.”

4.3.1.3.2 Sources of Referents

Examples of input validation referent, i.e., the source of the input data, in order of increasing validity, are SME opinion, output from another model, output from another validated/pedigreed model, data from a documented authoritative source such as the CRC Handbook of Chemistry and Physics, or experimental data fully representative of the system being modelled. Consider this aspect of the input when determining whether the referent meets threshold, objective, or neither.

Examples of verification testing frameworks in order of increasing relevance to an M&S that will be used in a fully integrated simulation environment are unit test, bench test, pair-wise integration test, and full integration test. These are some of the types of testing frameworks to consider when assessing the relevance of the verification testing framework to the M&S SIU framework.

Examples of results validation referents, i.e., the source of real-world comparison data, in order of increasing validity are SME opinion, bench test results, benchmark simulation results, real-world calibration test results, real-world test results from simplified scenarios, or real-world results from scenarios representative of the M&S SIU. Consider these types of sources when determining whether the referent meets threshold, objective, or neither.

4.3.1.3.3 Methodologies

Examples of input validation methodology, in order of increasing rigor, are methods such as visual examination of raw and transformed input data, manual spot checks of appropriate input loading, and correct mathematical transformation and automated checks of appropriate loading and transformation across the full input parameter space. Consider these aspects of and types of methodologies when determining whether the input validation methodology meets threshold, objective or neither.

Examples of requirement verification methodology, in order of increasing rigor, are code review only, code review plus visual results comparison against specification, code review plus quantitative results comparison against specification, and code review plus automated results comparison against independent code built to the same specification. These are some of the types of verification methodology to consider when assessing the rigor of the requirement verification.

Examples of M&S results validation methodology, in order of increasing rigor, are visual examination of M&S reconstruction results versus referent data without specific acceptability criteria, visual examination with acceptability criteria, mathematical comparison of results and referent without acceptability criteria, mathematical comparison with acceptability criteria, and full statistical comparison of results versus referent utilizing accepted statistical methods and confidence limits. Consider these types of analysis when considering whether the methodology meets objective, threshold, or neither with respect to the rigor of the analysis necessary to have confidence in the M&S results for the SIU.

4.3.2 E: Effects

Table 10 provides the workbook for the effects factor, E and Table 11 provides guidance for completing the factor E workbook.

Table 10: Workbook for Factor E.

E (Effects) Subfactor 'a': M&S Use (3=TRUE, 1=FALSE)					
Area	Atom	Statement	Value	Src	Comment
M&S Dependence	r	Supplemental Use			
	s	Secondary Use			
	t	Primary Use			
	u	Exclusive Use			
M&S Areas Addressed	v	Single Low Risk Area			
	w	Single Medium Risk Area			
	x	Single High Risk Area			
	y	Multiple Medium - Low Risk Areas			
	z	Multiple High Risk Areas			

Table 11: Guidance for Evaluation of Factor E.

Sub-Factor	Atom Statement (T/F)	Considerations
M&S Impact	Single low risk area	Answer is "True" if the requirement in question addresses a single area of the simulation space, and errors in that area would pose a low risk that the M&S will fail to meet the SIU. Otherwise, answer is "False."
	Single medium-risk area	Answer is "True" if the requirement in question addresses a single area of the simulation space, and errors in that area would pose a medium risk that the M&S will fail to meet the SIU. Otherwise, answer is "False."
	Single high risk area	Answer is "True" if the requirement in question addresses a single area of the simulation space, and errors in that area would pose a high risk that the M&S will fail to meet the SIU. Otherwise, answer is "False."
	Multiple low- or medium risk areas	Answer is "True" if the requirement addresses multiple areas of the simulation space, and requirement verification or validation failures would pose a low or medium risk that the M&S will fail to meet the SIU. Otherwise, answer is "False."
	Multiple high risk areas	Answer is "True" if the requirement addresses multiple areas of the simulation space, and requirement verification or validation failures would pose a high risk that the M&S will fail to meet the SIU. Otherwise, answer is "False."
M&S Reliance	Supplemental use	Answer is "True" if the M&S will be employed with other non-simulation methods to support the decision, development, or training process and will only provide additional confirmatory data for data already available through other means. Otherwise, answer is "False."
	Secondary use	Answer is "True" if the M&S will be employed with other non-simulation methods to support the decision, development, or training process and will provide significant data unavailable through other means. Otherwise, answer is "False."
	Primary use	Answer is "True" if the M&S will be employed with other non-simulation methods to support the decision, development, or training process, and it will provide the majority of the data required. Otherwise, answer is "False."
	Exclusive use	Answer is "True" if the M&S will be the only tool employed to produce data supporting the decision, development, or training process. Otherwise, answer is "False."

4.4 Outcomes of the MURM

4.4.1 Visualizing Results

The user of the workbook implementation enters True or False for each of the statements as it pertains to the requirement under consideration. Once all the evaluations have been entered, the workbook produces an M&S use risk value for the requirement, as shown in Table 12, and the user can plot the M&S use risk. Displaying the M&S use risk for multiple requirements simultaneously shows the analyst the distribution of M&S use risk and helps identify areas of the M&S that are especially vulnerable. The analyst can further enhance the visualization of the risks using techniques such as marker associations, as shown in Figure 12, where the marker is labelled with the requirement number and the color indicates the disposition (completeness) of the requirement. This additional information helps to quickly identify possibilities for M&S use risk balancing.

Table 12: Risk Constellation Detail.

Requirement #	p_C	p_E	M&S Use Risk
2	0.930	0.025	Very Low
4	0.225	0.735	Low
8	0.600	0.543	High

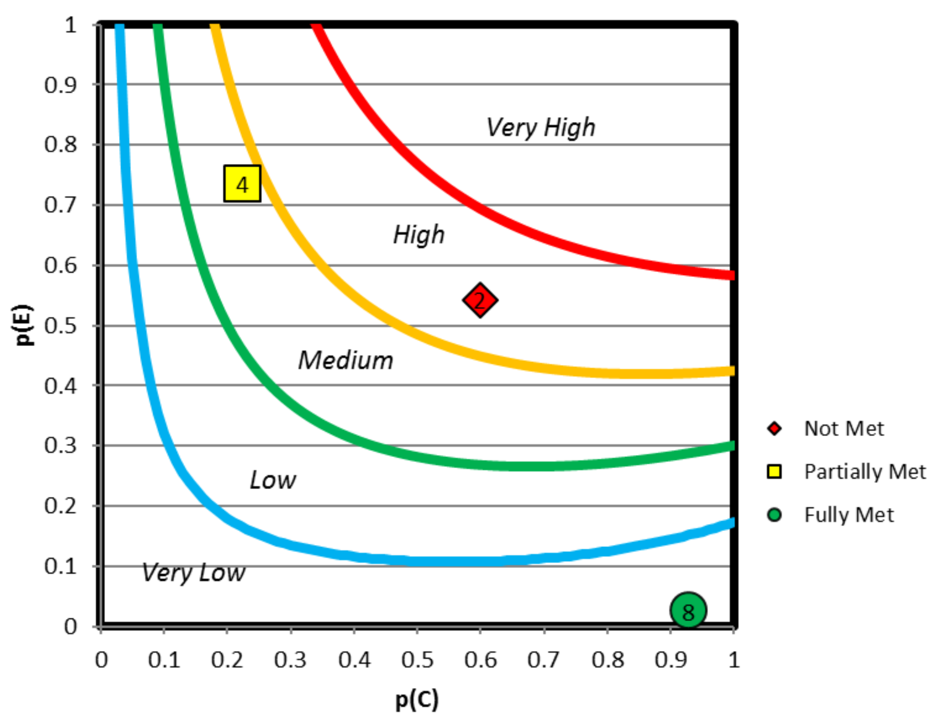

Figure 12: M&S Use Risk Plot.

4.4.2 Identifying Mitigation Strategies

Anytime the M&S use risk is evaluated, the result should be assessed with the stakeholders to determine whether the level of risk is acceptable. If the M&S use risk is accepted, the development can proceed as planned. If the M&S use risk is not accepted, then the V&V agent should work with the stakeholders through the process defined in Figure 13 to reduce the M&S use risk to an acceptable level.

First, the M&S SIU statement and detailed M&S requirements should be reviewed and refined if necessary, to ensure requirements are valid and testable. Then the acceptability criteria can be improved possibly by being quantified. Research can be conducted to identify sources of referent information. All these considerations can reduce the M&S use risk by directly impacting the evaluation of C1. The M&S limitations should be reviewed, and mitigation strategies defined or revised if possible, thereby changing the evaluation of C2. The V&V tasks and activities can be further tailored, or investment in more or better referent data can change the evaluation of C3. Finally, if the M&S use risk remains intolerable, the program or project should reconsider its test plan to improve the validation data available or reduce the reliance on the M&S with other types of system assessments.

As each activity is completed, the M&S use risk constellation will change. Therefore, within the sequence of activities, the V&V agent may recalculate the M&S use risk to determine the impact of each activity alone or in combination with other activities. This provides a sizeable number of options for reducing the M&S use risk to acceptable levels for the decision maker.

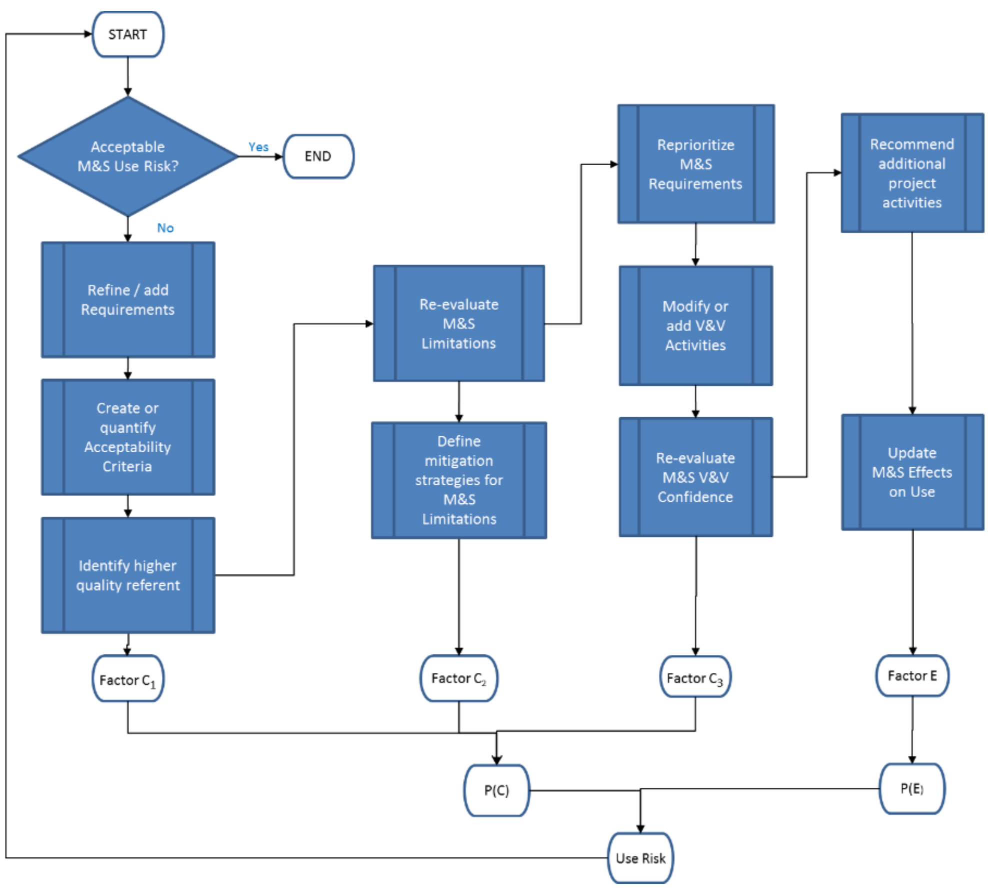

Figure 13: Process to Reduce Unacceptable Levels of M&S Use Risk.

5.0 USE CASE: SANDIA NATIONAL LABORATORIES CHALLENGE PROBLEM – TANKSIM

In 2015, JHU/APL collaborated with Dr. Kenneth Hu, the author of Sandia National Laboratories' (SNL's) 2014 V&V Challenge Problem [13], to demonstrate the MURM's use in a context familiar to the V&V community. Several academic and DoD organizations submitted responses to the 2014 V&V Challenge Problem, including a separate SNL team [14]. To demonstrate the use of the MURM, the SNL response was chosen because of the SNL team's superior scope and completeness in addressing the challenge.

Because the MURM is applied at the individual requirement level and designed for use with the requirements formalism typical of DoD M&S applications, JHU/APL collaborated with Dr. Hu to derive a SIU, HLRs, and DLRs from the Challenge Problem narrative. This led to organizing the requirements as illustrated in Figure 14. Note that for the sake of the M&S requirements identification scheme, the totality of the SNL team's response[2] was considered a single "model" arbitrarily named and hereafter referred to as TANKSIM.

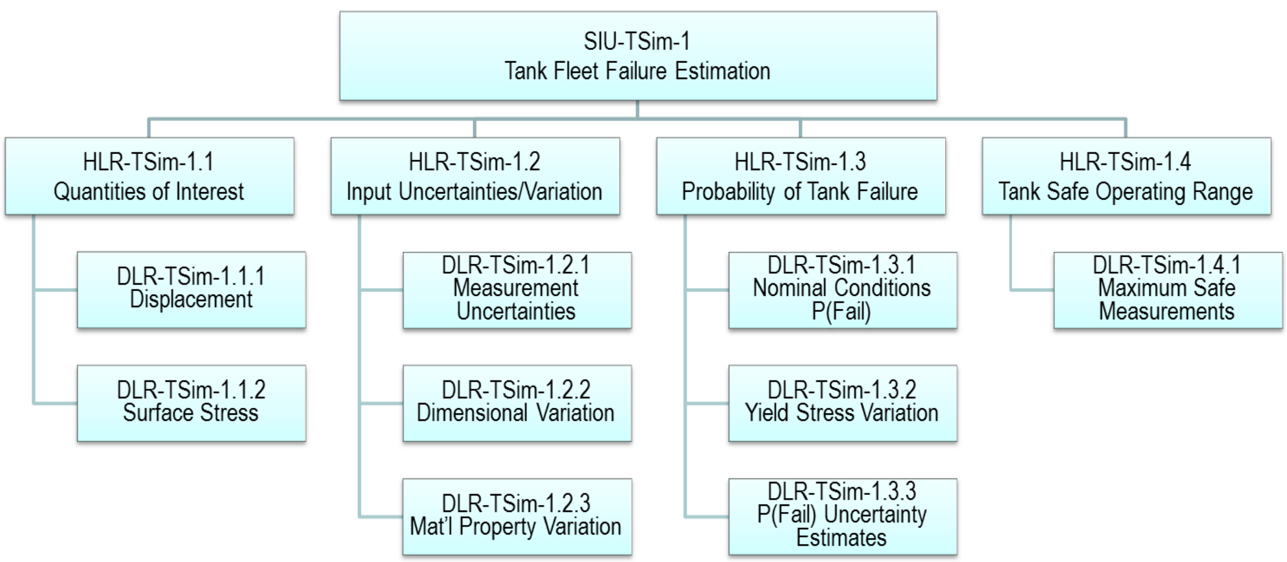

Figure 14: Structure of SIU, HLRs, and DLRs.

5.1 The TANKSIM Problem

The TANKSIM problem opened with the following back-story: The Mystery Liquid Company maintains a fleet of tanks that holds its mystery liquid under pressure. The pressure causes deformation of tank walls, and, during a routine safety inspection, the results of measurements on one tank produced an out-of-specification reading under normal loading conditions. This out-of-specification tank and two neighboring tanks were taken out of service and subjected to additional tests. Additionally, four tanks in different locations underwent limited tests while still in service. A modelling study was commissioned to supplement these tests, with the hypothesis being that the historic safety margin was being violated, and a better understanding of the margin to failure was therefore required. The ultimate goal was to decide whether the remaining tanks may remain in service or if they must be retired immediately.

[2] The SNL team utilized a finite difference code and manual pre and post processing calculations to perform the modelling analysis.

The following data were provided in the challenge problem statement:

- The mystery liquid's known relationship between its composition and specific weight;
- Laboratory results for the out-of-specification tank (Tank 0) and its two neighbors (T1 and T2) such as Young's modulus, Poisson's ratio, and yield stresses;
- Field test results for four tanks (T3, T4, T5, and T6), such as radii and lengths;
- Manufacturer original specifications for the tanks (dimensions and material properties);
- Computations for a model predicting stress and displacements for pressurized tanks (T1 and T2) to compare with actual displacement data; and
- Field measurement displacements for Tanks T3 through T6 when under various pressures and liquid loadings.

A simplified finite-difference model was used to predict failure at extreme pressures and liquid loadings. The model relates the displacement (d) and surface Von Mises stress (σ) to the following:

- Tank dimensions: radius (R), length (L), thickness (T);
- Material properties: Young's modulus (E), Poisson's ratio (ν);
- Operating conditions: pressure (P), liquid height (H);
- Liquid property: specific gravity (χ); and
- Model parameter: mesh size (m).

As a stipulation of the problem, code verification was not possible. The model also had some limitations; the most egregious was that it modelled only the cylindrical section of the tank and not the end caps.

The tank is predicted to fail if the model computed a Von Mises stress anywhere on the tank surface that exceeds the wall material's yield strength. The challenge of the problem was to execute an analysis strategy to predict failure probabilities for two scenarios, followed by an assessment of the credibility of the predictions and a recommendation of whether (or not) to retire the tanks.

5.2 MURM Evaluation of the SNL Solution

In the context of this challenge, the MURM is applied not to offer a solution to the challenge problem but to assess the risk involved in using the recommendation resulting from SNL's use of the provided model, the pre- and post-processing, and the provided problem information.

The risk in using the model's recommendation is approached by evaluating the M&S use risk associated with each of TANKSIM's DLRs. Detailed descriptions of the nine TANKSIM DLRs are contained in Table 13.and Table 14 details the factor evaluations for DLR-TS-1.2.3 as an example of the input from the V&V agent. The M&S use risk for each DLR is illustrated in Figure 15. The results of the TANKSIM MURM baseline assessment indicate that seven of the nine DLRs are in the high to very high region of M&S use risk. This means that there is currently a substantial risk of inappropriate application of TANKSIM results for the SIU, SIU-TSim-1, producing unacceptable consequences to the decision maker. This level of TANKSIM use risk may not be tolerable to the decision maker for making a recommendation, but this is the baseline assessment based on the data presented in the problem statement and the SNL solution.

Table 13: Description of TANKSIM DLRs.

REQ ID	Name	Description
SIU-TSim-1	Tank Fleet Failure Estimation	The TANKSIM (TSim) M&S will be used to determine whether the fleet of Mystery Liquid storage tanks have a low enough Probability of Failure that they can be kept in service for "a few years" while replacements are ordered. Specifically, the M&S will be used to make a decision on whether to remove all tanks from service [immediately] or modify operating limits [until new tanks are installed and operational].
HLR-TSim-1.1	Quantities of Interest	TANKSIM shall produce certain quantities of interest (QoIs).
DLR-TSim-1.1.1	Displacement	TANKSIM shall produce displacement normal to the tank surface d = (x, ϕ, P, H, χ, E,ν, L, R, T).
DLR-TSim-1.1.2	Surface Stress	TANKSIM shall produce surface Von Mises stress σsurf = (x,ϕ, P, H, χ, E,ν, L, R, T).
HLR-TSim-1.2	Input Uncertainties/ Variation	TANKSIM shall be capable of characterizing QoI variation due to input uncertainties.
DLR-TSim-1.2.1	Measurement Uncertainties	TANKSIM shall be capable of characterizing d and σsurf variation due to measurement uncertainties for P, H, and χ.
DLR-TSim-1.2.2	Dimensional Variation	TANKSIM shall be capable of characterizing d and σsurf variation due to tank dimensional uncertainties for L,R, and T.
DLR-TSim-1.2.3	Material Property Variation	TANKSIM shall be capable of characterizing d and σsurf variation due to material property uncertainties for E and ν.
HLR-TSim-1.3	Probability of Tank Failure	TANKSIM shall calculate the [current] probability of tank failure.
DLR-TSim-1.3.1	Nominal Conditions P(Fail)	TANKSIM shall calculate P(Fail)l for the nominal test condition specified in the Problem Statement: P = 73.5 psig, χ = 1, H = 50in. A tank will be considered failed if σsurf > σyield for any location (x, ϕ) on the tank walls.
DLR-TSim-1.3.2	Yield Stress Variation	TANKSIM shall be capable of characterizing variation in yield stress σyield.
DLR-TSim-1.3.3	P(Fail) Uncertainty Estimates	TANKSIM shall produce uncertainty estimates for P(Fail).
HLR-TSim-1.4	Tank Safe Operating Range	TANKSIM shall be capable of calculating the range of "safe" operating condition measurements.
DLR-TSim-1.4.1	Maximum Safe Measurements	TANKSIM shall be capable of calculating the maximum measured P, H, χ such that P(Fail) < 10-3.

Table 14: MURM Evaluation of DLR-TS-1.2.3.

ID	DLR-TS-1.2.3				
Title	Material Property Variation				
Requirement Text	TANKSIM shall be capable of characterizing material property uncertainties for *E* and v.				
Acceptance Criteria					
C1 (Clarity of SIU): Sub-Factor 'a' Requirements (3 = TRUE, 1 = FALSE)					
Area	**Atom**	**Statement**	**Value**	**Src**	**Comment**
Requirements	z	Quantitative basis for Requirement established	3	This document	Quantitative scalars
C1 (Clarity of SIU): Sub-Factor 'b' Acceptability Criteria (3 = TRUE, 1 = FALSE)					
Area	**Atom**	**Statement**	**Value**	**Src**	**Comment**
Acceptability Criteria	y	Acceptability Criteria exists	1	This document	See above
	z	Quantitative basis for Acceptability Criteria established	1	This document	See above
C1 (Clarity of SIU): Sub-Factor 'c' Verification Referent (3 = TRUE, 1 = FALSE)					
Area	Atom	Statement	Value	Src	Comment
Verification Referent	x	Verification Referent Identified	3	SNL Personnel verbal	Verification Test Suite
	y	Verification Referent Source established	3	SNL Personnel verbal	Verification Test Suite already exists
	z	Verification Quantitative basis for referent established	3	SNL Personnel verbal	Verification Test Suite is quantitative in nature.
C1 (Clarity of SIU): Sub-Factor 'd' M&S Results Validation Referent (3 = TRUE, 1 = FALSE)					
Area	**Atom**	**Statement**	**Value**	**Src**	**Comment**
Validation Referent	x	M&S Results Validation Referent Identified	3		Results validation N/A for this requirement. Set to true per MURM methodology.

C1 (Clarity of SIU): Sub-Factor 'd' M&S Results Validation Referent (3 = TRUE, 1 = FALSE) (cont'd)					
Area	**Atom**	**Statement**	**Value**	**Src**	**Comment**
	y	M&S Results Validation Referent Source established	3		Results validation N/A for this requirement. Set to true per MURM methodology.
	z	Validation Quantitative basis for referent established	3		Results validation N/A for this requirement. Set to true per MURM methodology.

C2 (M&S Limitations) (3 = TRUE, 1 = FALSE)					
Area	**Atom**	**Statement**	**Value**	**Src**	**Comment**
Mitigation	s	Current mitigation status of all known limitations pertinent to the requirement is 'not required'.	1	DRAFT-VVUQ-15-1009-1	Except for L, calibration against displacement data produced confidence intervals that were unrealistically large and encompassed parameter values that are not physically meaningful. Mitigation strategy was to fall back on confidence intervals directly generated by Minitab analysis of the admittedly sparse experimental data.
	t	Current mitigation status of all known limitations pertinent to the requirement is 'complete'.	3		
	u	Current mitigation status of all known limitations pertinent to the requirement is 'partially complete'.	1		
	v	Current mitigation status of all known limitations pertinent to the requirement is 'impossible'.	1		
Consequences	w	Consequences to the SIU of all known M&S limitations pertinent to the requirement are negligible.	1	DRAFT-VVUQ-15-1009-1	SNL team noted that the extremely large confidence intervals and the discrepancy between calibrated and experimental material relationships raise concerns regarding the credibility of the ensuing analysis.

C2 (M&S Limitations) (3 = TRUE, 1 = FALSE) (cont'd)

Area	Atom	Statement	Value	Src	Comment
Consequences (cont'd)	x	Consequences to the SIU of all known M&S limitations pertinent to the requirement are minor.	1		In short, they had no pedigree or uncertainty information for the materials data. As a result, they cannot determine if the data is adequate for calibration needs. Without the ability to establish credibility of the data, we could not assess to what extent model form uncertainty may be responsible for their observations. This would seem to have serious consequences for this requirement.
	y	Consequences to the SIU of all known M&S limitations pertinent to the requirement are serious.	3		
	z	Consequences to the SIU of all known M&S limitations pertinent to the requirement are grave.	1		

C3 (Confidence) Sub-Factor 'a': Input Validation Analysis (3 = TRUE, 1 = FALSE)

Area	Atom	Statement	Value	Src	Comment
Scope	u	Scope Meets Objective	3	DRAFT-VVUQ-15-1009-1	The extent and relevance of the confidence interval analysis are both full. Meets objective.
	v	Scope Meets Threshold	3		
Referent	w	Referent Meets Objective	1	DRAFT-VVUQ-15-1009-1	Experimental data used to calculate confidence intervals was sparse. Meets threshold but not objective.
	x	Referent Meets Threshold	3		
Methodology	y	Methodology Meets Objective	1	DRAFT-VVUQ-15-1009-1	Minitab confidence interval calculations were meant to be confirmed or supplanted by intervals from calibration but that didn't work out. Meets threshold but not objective.
	z	Methodology Meets Threshold	3		

C3 (Confidence) Sub-Factor 'b': Verification Analysis (3 = TRUE, 1 = FALSE)					
Area	**Atom**	**Statement**	**Value**	**Src**	**Comment**
Scope	u	Scope Meets Objective	3	DRAFT-VVUQ-15-1009-1	E and v were varied in the sensitivity analysis so that provided a full extent to the verification that TANKSIM can account for uncertainties in these parameters. The sensitivity analysis was performed on the same version of TANKSIM as would be used for the SIU, so the relevance is full. Full/full = Meets objective.
	v	Scope Meets Threshold	3		
Framework	w	Framework Meets Objective	3	DRAFT-VVUQ-15-1009-1	The sensitivity analysis was performed on a standalone version of TANKSIM just like the SIU would be, so the verification framework meets objective.
	x	Framework Meets Threshold	3		
Methodology	y	Methodology Meets Objective	3	DRAFT-VVUQ-15-1009-1	To the best of my knowledge, the verification of this requirement via the sensitivity analysis and calibration analysis is sufficient for the methodology assessment to be 'meets objective'.
	z	Methodology Meets Threshold	3		
C3 (Confidence) Sub-Factor 'c': M&S Results Validation Analysis (3 = TRUE, 1 = FALSE)					
Area	**Atom**	**Statement**	**Value**	**Src**	**Comment**
Scope	u	Scope Meets Objective	3	DRAFT-VVUQ-15-1009-1	Results validation N/A for this requirement. With no need for results validation, scope meets objective per DoD convention.
	v	Scope Meets Threshold	3		

C3 (Confidence) Sub-Factor 'c': M&S Results Validation Analysis (3 = TRUE, 1 = FALSE) (cont'd)					
Area	**Atom**	**Statement**	**Value**	**Src**	**Comment**
Referent	w	Referent Meets Objective	3	DRAFT-VVUQ-15-1009-1	Results validation N/A for this requirement. With no need for results validation, referent meets objective per DoD convention.
	x	Referent Meets Threshold	3		
Methodology	y	Methodology Meets Objective	3	DRAFT-VVUQ-15-1009-1	Results validation N/A for this requirement. With no need for results validation, methodology meets objective per DoD convention.
	z	Methodology Meets Threshold	3		

E (Effects) Sub-Factor 'a': M&S Use (3 = TRUE, 1 = FALSE)					
Area	**Atom**	**Statement**	**Value**	**Src**	**Comment**
M&S Dependence	r	Supplemental Use	1	SAND2013-10486P	Experimental data to characterize uncertainty in these parameters is available, but TANKSIM is the primary means for characterizing the uncertainty for the SIU of predicting tank failure.
	s	Secondary Use	1		
	t	Primary Use	3		
	u	Exclusive Use	1		
M&S Areas Addressed	v	Single Low Impact Area	1	SAND2013-10486P	Multiple parameters in this requirement means multiple areas addressed. The relatively lower partial rank correlations for these parameters, determined by the sensitivity analysis, suggest medium risk to the overall SIU.
	w	Single Medium Impact Area	1		
	x	Single High Impact Area	1		
	y	Multiple Medium - Low Impact Areas	3		
	z	Multiple High Impact Areas	1		

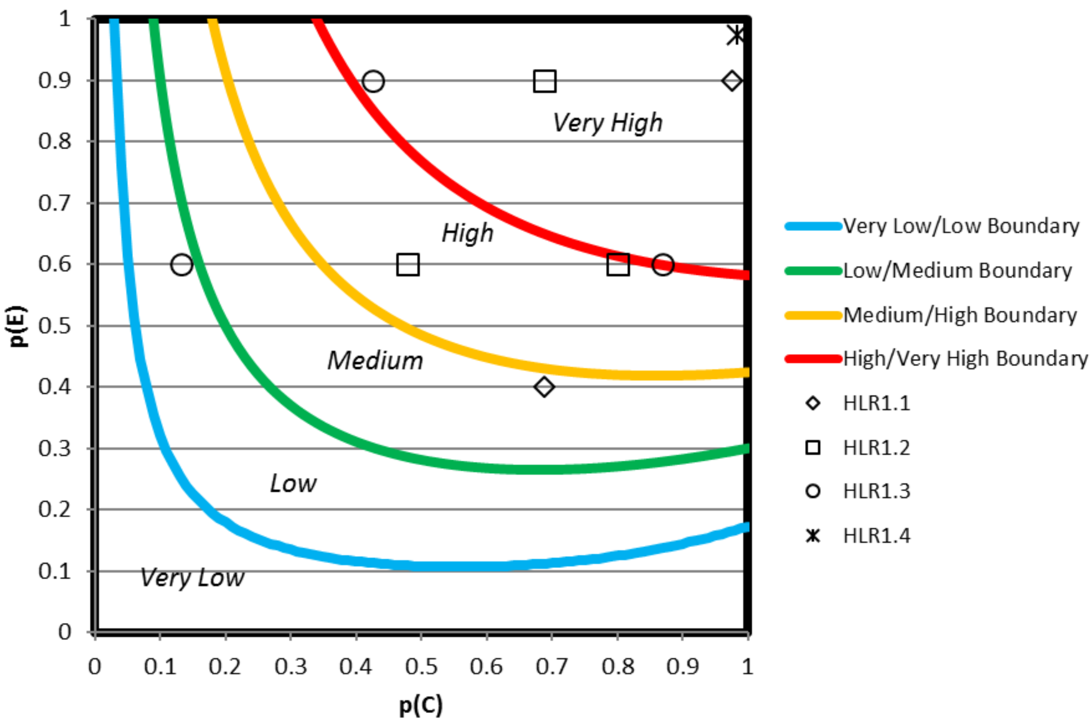

Figure 15: TANKSIM Baseline Assessment.

The MURM's usefulness beyond the initial baseline assessment comes by examining the detailed answers to the MURM sub-factors and atoms for each DLR to determine potential remedial activities and assessing the impact of those activities (i.e., changing "no" to "yes" for one or more atoms) on the M&S use risk. This predictive utilization of the MURM can then be used to determine which remedial actions offer the largest reduction in M&S use risk.

The recommended remedial actions for TANKSIM and its cumulative impacts are as follows:

1) **Add quantitative acceptability criteria to all requirements**. This action improves clarity of SIU and affects all DLRs. Figure 16 illustrates the movement of the risk constellation resulting from this action.
2) **Perform more tank measurements**. Complete more coupon tests at critical locations and perform more tank dimensional measurements across the tank fleet to better characterize variation in tank fleet material and dimensional properties. This action affects all DLRs to varying degrees except DLR-TSim-1.2.1. Figure 17 illustrates the cumulative improvement of this action in addition to that of action 1.
3) **Address gauge uncertainties**. Complete tests of representative samples of measurement devices to better characterize measurement uncertainties across the tank fleet and incorporate gauge uncertainty into the model analysis. This action affects DLR-TSim-1.2.1. Figure 18 illustrates the cumulative improvement of this additional action to those of actions 1 and 2.
4) **Improve model representation of tank geometries**. Use a full finite element model to more closely represent the actual tank geometry (e.g., hemispherical ends, legs, bracing). This action affects DLR-TSim-1.1.1 and DLR-TSim-1.1.2. The cumulative effects of actions 1 through 3 and this action are shown in Figure 19.

5) **Perform more destructive tank testing**. Perform destructive tank testing on multiple tanks to determine failure loading conditions and probability of failure. Then validate modelled P(fail) against actual failure data. This action affects DLR-TSim-1.3.1. Figure 20 shows the cumulative effects of actions 1 through 4 and this action. Note that this action reduces both the p_C and p_E values for this DLR, as emphasized by the arrow in the figure. Typically, remedial actions only improve the MURM C-factors and lowering the p_C score, but in this case, performing destructive tank testing also improves the E factor by reducing this DLR's reliance on modelling by providing additional test-derived data for estimating P(fail).

6) **Address P(fail) uncertainty estimation**. Examine the P(fail) uncertainty estimation requirement specified in the challenge problem statement and perform all aspects of V&V related to this requirement This action affects DLR-TSim-1.3.3. The cumulative effects of 1 through 5 and this action are shown in Figure 21.

7) **Address tank safe operating range question**. Examine the requirement to determine the maximum nominal loading conditions for safe tank operation [i.e., P(fail) < 10^{-3}], and perform all aspects of V&V related to this requirement. For reasons similar to DLR-TSim-1.3.3, this requirement was not addressed by the SNL team. This action affects DLR1.4.1. The cumulative effects of actions 1 through 6 and this action are shown in Figure 22.

Figure 22 shows the cumulative effects of all of the proposed remedial actions. M&S use risks for six of the nine DLRs have moved to the low risk region, and the remaining three are in the medium risk region. Under these circumstances, a decision maker could be much more confident in a decision based on TANKSIM results (i.e., whether to continue with current tank usage or immediately replace) than can be made with the baseline assessment.

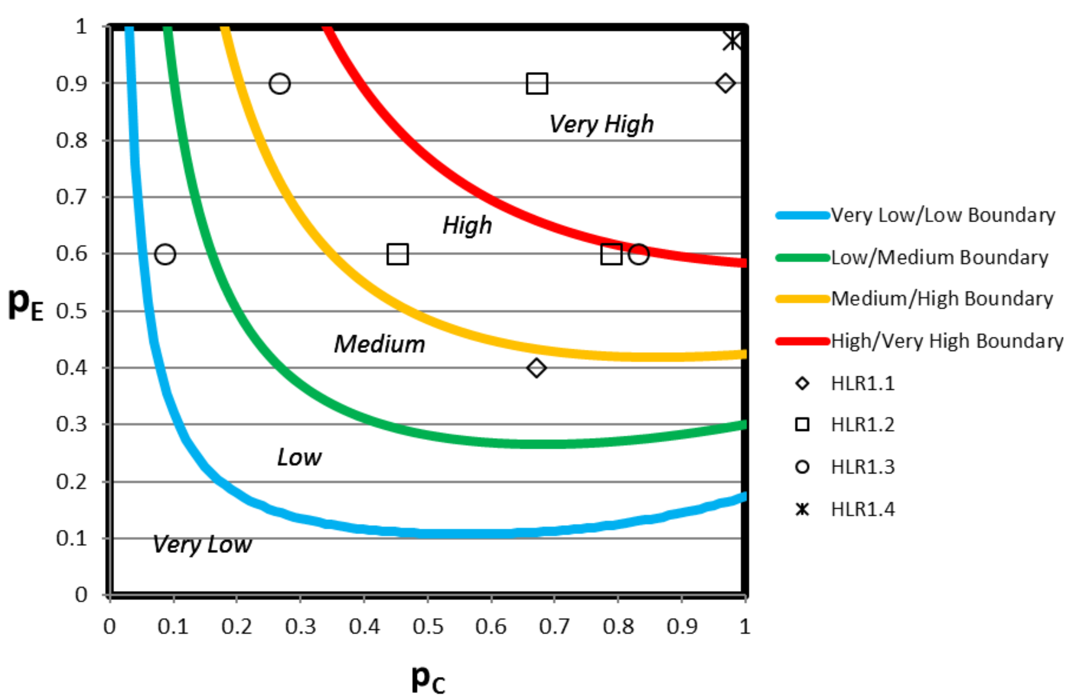

Figure 16: Add Quantitative Acceptability Criteria.

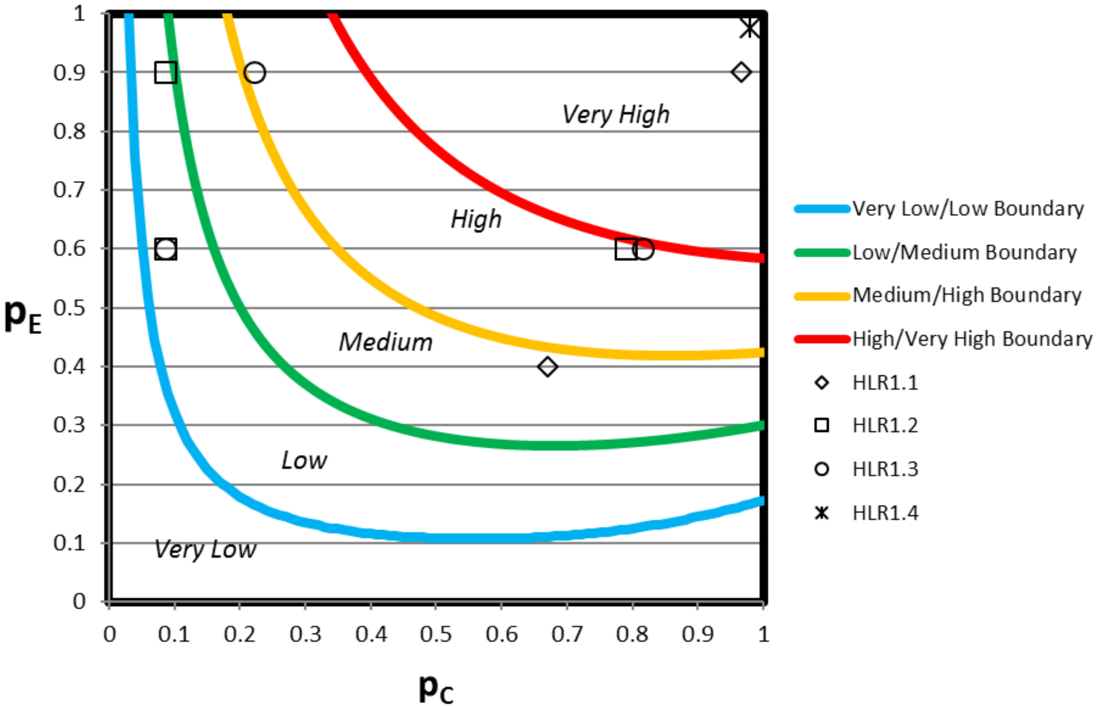

Figure 17: Plus Perform More Tank Measurements.

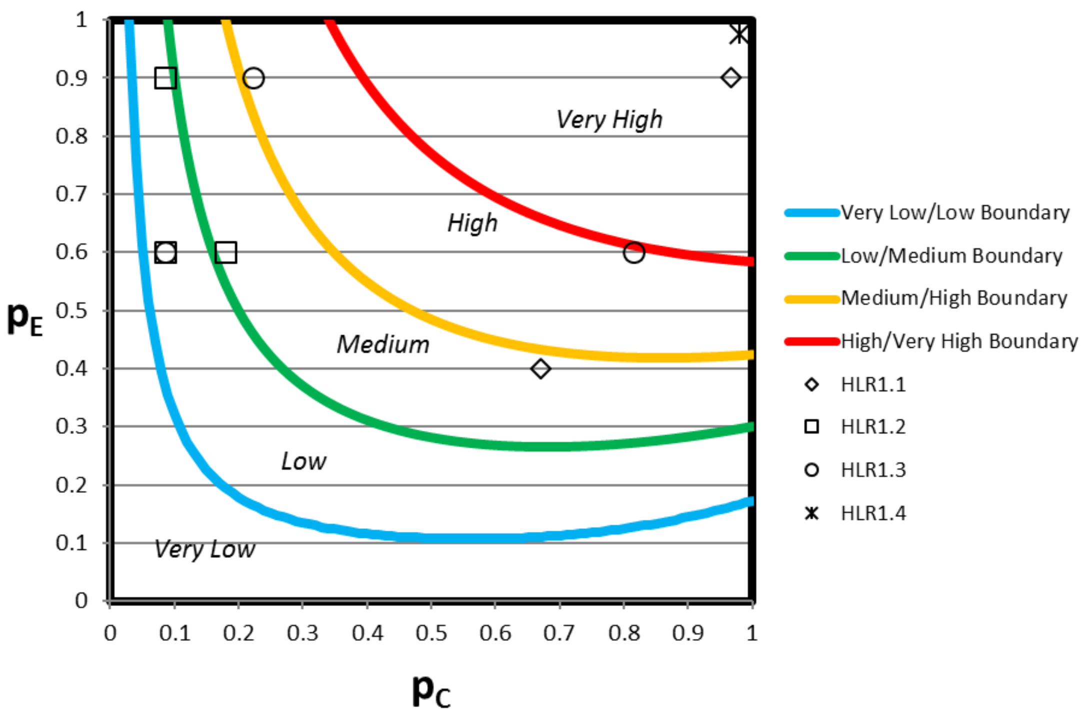

Figure 18: Plus Address Gauge Uncertainties.

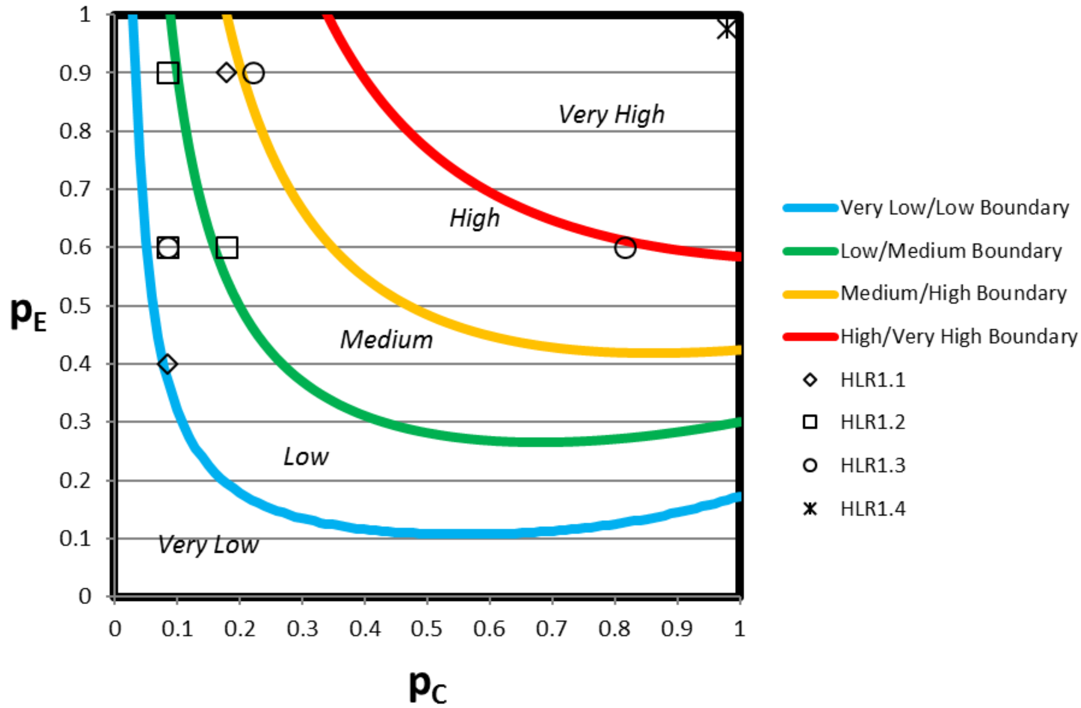

Figure 19: Plus Representative Tank Geometry.

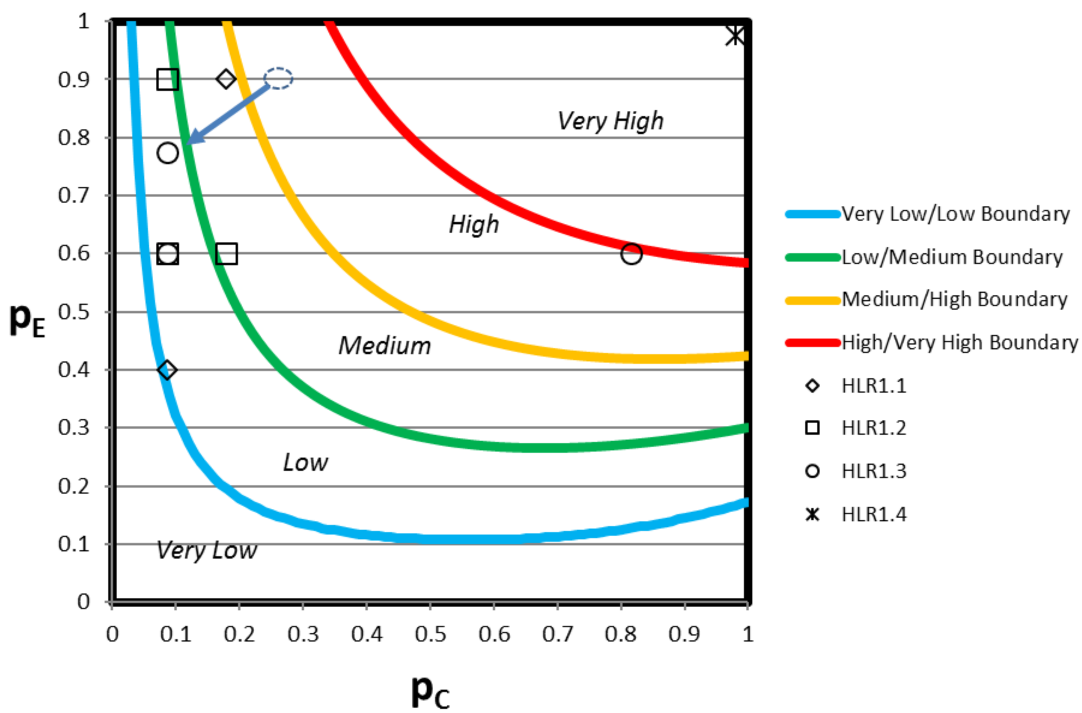

Figure 20: Plus Data from Destructive Tank Testing.

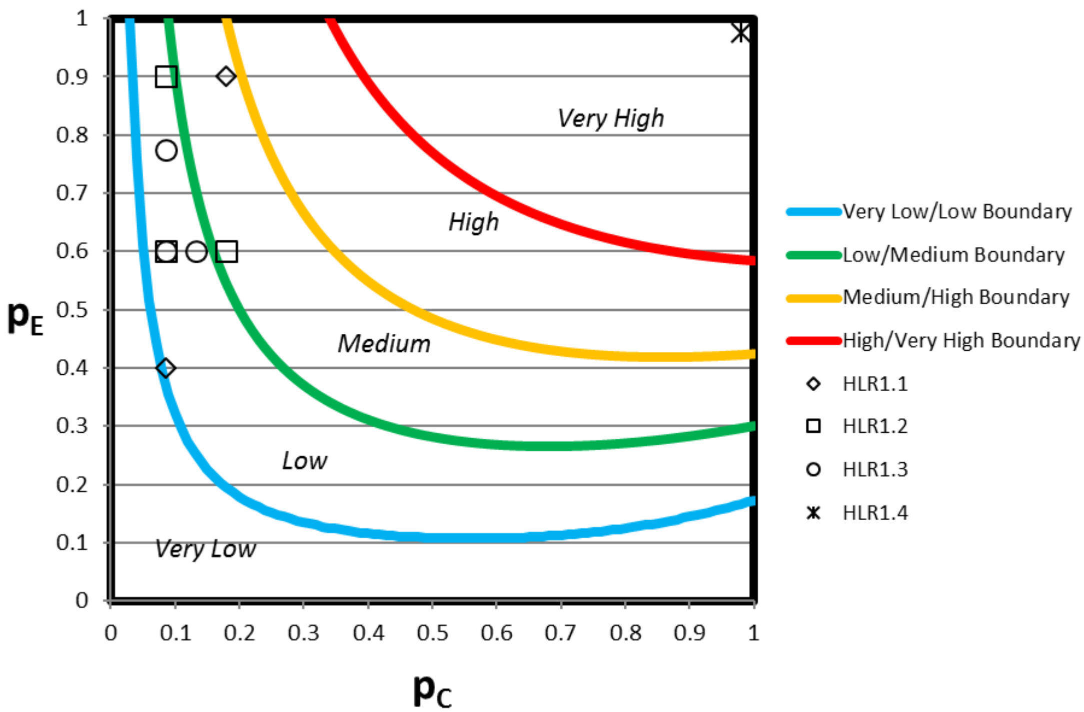

Figure 21: Plus Address P(fail) Uncertainty Estimation.

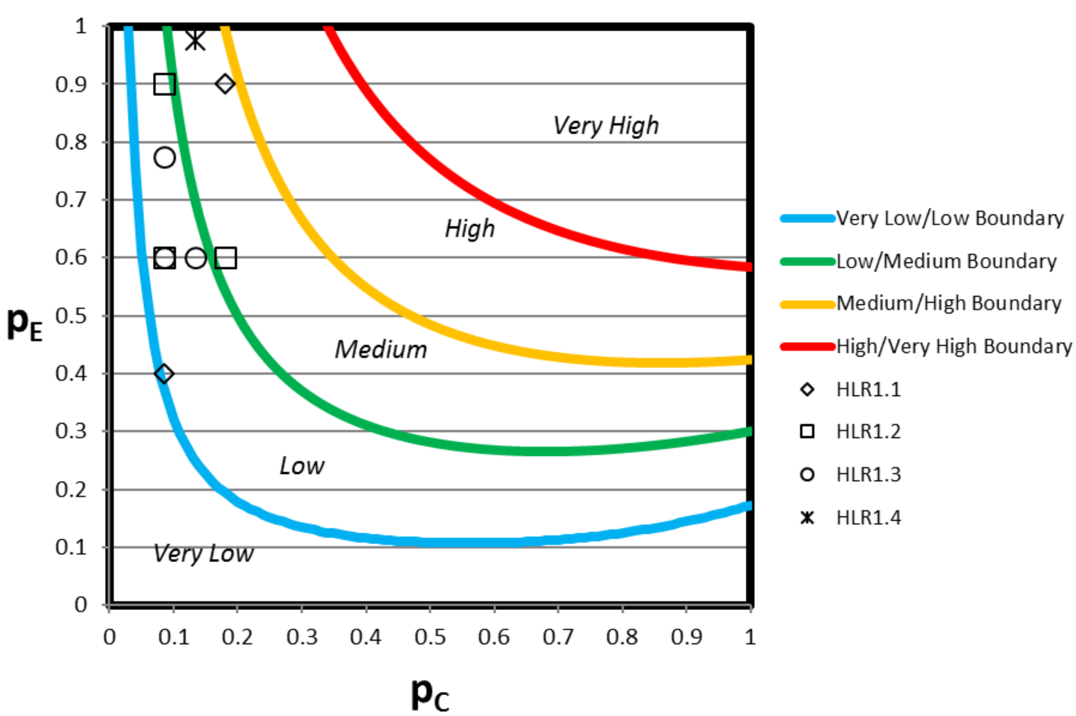

Figure 22: Plus Address Maximum Safe Loading.

6.0 CONCLUSIONS AND RECOMMENDATIONS

6.1 Conclusions

MSG-139 satisfied the objectives set out in the Terms of Reference. Using a thorough scientific approach, the group:

- Reviewed the current state of M&S use risk assessment;
- Evaluated available methodologies;
- Selected the MURM based on viability and applicability of the MURM to a broad set of problems; and
- Matured and enhanced the MURM to a state useful to the community.

In the course of completing its work, the task group codified a methodology to plan, identify, and mitigate M&S use risk. Through a series of discussions and examination of various uses of M&S in their respective domains, the task group members concluded that the MURM is applicable across the M&S development life cycle. Specifically, evaluation of M&S use risk at the outset of an M&S development project based on HLRs is useful for planning M&S development, prioritizing function development, and tailoring V&V Plans. Once M&S development begins and external forces cause deviations from the development and V&V Plans, the MURM can be used to assess and mitigate the impact of those changes. Finally, as the M&S development nears completion and the V&V activities end, the MURM provides critical information to the M&S user.

Finally, the task group developed and used educational materials for the workshop conducted to facilitate development of additional use cases by the task group participants. The workshop also provided a forum to assess and improve the education materials.

6.2 Recommendations

To increase confidence in the methodology, build consensus regarding breadth of its applicability, and contribute to the NATO body of knowledge before widespread dissemination of the implementation and related documentation, the task group recommends the following activities:

- Develop more use cases to increase the experience base, potentially improve the methodology, and further develop visualization strategies.
- Develop and provide continuing education to the M&S community of practice through workshops and lecture series. The MURM workshop materials represent a significant step toward the materials needed for such activities.
- Develop an international standard subject to configuration management and change control through the SISO using its Product Development Group process.

7.0 REFERENCES

[1] NATO RTO: RTO Task Group, M&S Use Risk Identification and Management Terms of Reference, 8 October 2014.

[2] JHU/APL, Risk Based Methodology for Verification, Validation, and Accreditation (VV&A) M&S Use Risk Methodology (MURM), JHU/APL Technical Report NSAD-R-11-011, April 2011.

[3] Simulation Interoperability Standards Organization (SISO), Guide for Generic Methodology for Verification and Validation (GM-VV) to Support Acceptance of Models, Simulations, and Data, SISO-GUIDE-001.1-2012, 5 October 2012.

[4] JHU/APL, An Approach for Realizing a Risk-Based VV&A (RBA) Methodology Comments Summary, JHU/APL Technical Report NSAD-R-11-018, January 2011.

[5] IEEE Computer Society, IEEE Recommended Practice for Verification, Validation, and Accreditation of a Federation, an Overlay to the High Level Architecture Federation Development and Execution Process, IEEE Std 1516.4™-2007, September 2007.

[6] Research and Technology Organization (RTO), Risk-Based Tailoring of the Verification, Validation, and Accreditation/Acceptance Processes, Final Report of Modelling and Simulation, Group 054 / Task Group 037, RTO Technical Report TR-MSG-054, March 2010.

[7] JHU/APL, The Risk-Based Methodology for Verification, Validation, and Accreditation (VV&A) Literature Survey and Methods Bibliography, JHU/APL Technical Report NSAD-R-09-207, 2009.

[8] JHU/APL, The Risk-Based Methodology for Verification, Validation, and Accreditation (VV&A) Assessment of Applicable Methods, JHU/APL Technical Report NSAD-R-10-001, 2010.

[9] JHU/APL, An Approach for Realizing a Risk-Based VV&A (RBA) Methodology, JHU/APL Technical Report NSAD-R-10-020, 2010.

[10] JHU/APL, Risk-Based Methodology for Verification, Validation, and Accreditation (VV&A): Literature Survey Update, JHU/APL Technical Report NSAD-R-18-024, March 2018.

[11] Department of Defense, Documentation of Verification, Validation, and Accreditation of Models and Simulations, MIL-STD-3022, January 2008.

[12] Harmon, S.Y., and Youngblood, S.M. Lessons Learned about Developing Acceptable Acceptability Criteria, Proceedings of the 2008 European Simulation Interoperability *Workshop,* paper 08E-SIW-056, pp. 16-19 June 2008.

[13] Hu, K.T., 2014 V&V Challenge: Problem Statement, Sandia National Laboratories SAND2013-10486P, 2013.

[14] Beghini, L.L. and Hough, P.D. Sandia V&V Challenge Problem: A PCMM-Based Approach to Assessing Prediction Credibility, Sandia National Laboratories VVUQ-15-1009-1, 2015.

[15] Jaynes, E.T., Probability Theory: The Logic of Science. New York: Cambridge University Press, 2003.

Appendix 1: DERIVATION OF M&S USE RISK EQUATION

The purpose of this appendix is to present the formal derivation of the M&S use risk equation introduced in Section 3.1.

The mathematical expressions contained herein (see Table A1-1) generally follow mathematical convention with the exception of probabilities of complex expressions. Probabilities of simple expressions will be shown as lowercase p with a subscript indicating the expression to which the probability applies (e.g., p_1). Probabilities of complex expressions will be shown as uppercase P with the expression to which the probability applies described in parentheses [e.g., $P(A \wedge B)$]. This approach facilitates reading of the document. The following notation is used throughout.

Table A1-1: Mathematical Expressions.

Symbol	Meaning
$\wedge$	and
$\vee$	or
$\neg$	not
$\Rightarrow$	implies

The statement of the M&S use risk equation was introduced in Section 3.1. This appendix provides the detailed mathematical derivation of that equation.

Recall the semantic definition of the M&S use risk:

> The probability that inappropriate application of M&S results for the intended use will produce unacceptable consequences to the decision maker.

Further recall the set of actors and operators inherent in the semantic definition from Section 1.6:

1) C is the inappropriate application of the M&S;
2) E is the manifestation of unacceptable consequences; and
3) The probability that there is a causal relationship between them.

The goal is to derive an expression for the M&S use risk in terms of the two known quantities, the prior probability of C, and the prior probability of E. This is accomplished in a series of logical steps:

1) Decompose the problem into two conditions.
2) Define the probability of the occurrence of each condition.
3) Define the possible states for the C/E space.
4) Define the probabilities associated with the states.
5) Apply constraints to the definition of the conditions to express them in terms of the probabilities of the states of C and E.

1) Decompose the Problem into Two Conditions

The M&S use risk is the coincidence of the following two conditions:

- There are instances of inappropriate use of the M&S (C) that cause instances of unacceptable consequences to the decision maker (E).
- The inappropriate use of the M&S (C) coincides with the unacceptable consequences to the decision maker (E).

Both conditions are required because not all coincidental conditions are related. It cannot be assumed that C and E are related just because they occur together, thereby avoiding the fallacy of *cum hoc ergo proctor hoc*. The logical implication that C causes E can only be true if both C and E occur. However, if C and E occur, a causal connection may or may not exist. The formulation of M&S use risk seeks to assign a probability to a causal relationship between the inappropriate use and the manifestation of the undesirable consequences to the decision maker. The existence of this connection may depend on a long series of experiments and/or analyses; however, that information may not be available at the outset of the M&S use risk analysis.

2) Define the Probability of the Occurrence of Each Condition

In keeping with the initial state of ignorance with respect to a relationship, the probability of a causal relationship between C and E is expressed as the probability of the canonical form of the implication [15]; that is, the implication is the disjunction "not C or E":

$$P(C \Rightarrow E) \equiv P(\neg C \vee E)$$

using the sum rule for evaluating probability:

$$P(\neg C \vee E) = P(\neg C) + p_E - P(\neg C)P(E|\neg C)$$

and substituting $P(\neg C) = 1 - p_C$,

$$P(\neg C \vee E) = 1 - p_C + p_E - (1 - p_c)P(E|\neg C)$$

For convenience, let $z = P(E|\neg C)$ so that:

$$P(\neg C \vee E) = 1 - p_C + p_E - (1 - p_c)z$$

The probability that C and E are coincident is equal to:

$$P(C \wedge E) = p_C \wedge p_E = p_C p_E$$

$$M\&S\ Use\ Risk = P(\neg C \vee E) \wedge P(C \wedge E)$$

3) Define the Possible States for the C-E Space

Table A1-2 shows the combinations of possible states for C and E. Namely:

- p_1 is the probability that there are instances of inappropriate use of the M&S and there are instances of unacceptable consequences of decisions.

- p_2 is the probability that there are instances of inappropriate use of the M&S and there are no instances of unacceptable consequences of decisions.
- p_3 is the probability that there are no instances of inappropriate use of the M&S and there are instances of unacceptable consequences of decisions.
- p_4 is the probability that there are no instances of inappropriate use of the M&S and there are no instances of unacceptable consequences of decisions.

Table A1-2: States for C and E.

Prior Information	State	C	E	P(State)
p_C, p_E	1	T	T	p1
	2	T	F	p2
	3	F	T	p3
	4	F	F	p4

4) Define the Probabilities Associated with the States

Inspection of the table yields:

$$p_1 + p_2 = p_C$$

$$p_1 + p_3 = p_E$$

Because the four states represent the entire probability space:

$$p_1 + p_2 + p_3 + p_4 = 1$$

$$z = P(E|\neg C) = \frac{p_3}{p_3 + p_4}$$

Thus, the solution for the state probabilities in terms of p_C, p_E, and z is the following:

$$p_1 = p_E - (1 - p_C)z$$

$$p_2 = p_C - p_E + (1 - p_C)z$$

$$p_3 = (1 - p_C)z$$

$$p_4 = (1 - p_C)(1 - z)$$

5) Apply Constraints to the Definition of the Conditions to Express Them in Terms of the Probabilities of the States of C and E.

The final constraint that is placed on the system that enables the expression of z, in terms of the known quantities p_C and p_E, is one of maximum entropy of the C-E space.

The information entropy of the C and E space, H, is given by:

$$H = -p_1 \log_2 p_1 - p_2 \log_2 p_2 - p_3 \log_2 p_3 - p_4 \log_2 p_4$$

To maximize the information entropy:

$$\frac{\partial H}{\partial z} = 0 \text{ and } \frac{\partial^2 H}{\partial z^2} < 0$$

After differentiation and simplification, the expression for the derivative is:

$$\frac{\partial H}{\partial z} = (1 - p_C) \log_2 \left[\frac{(p_E - (1 - p_C)z)(1 - p_C)(1 - z)}{(p_C - p_E + (1 + p_C)z)(1 - p_C)z} \right]$$

To maximize:

$$\frac{\partial H}{\partial z} = 0$$

Thus, either:

$$(1 - p_C) = 0; p_C = 1$$

Or:

$$\log_2 \left[\frac{(p_E - (1 - p_C)z)(1 - p_C)(1 - z)}{(p_C - p_E + (1 + p_C)z)(1 - p_C)z} \right] = 0$$

$$\frac{(p_E - (1 - p_C)z)(1 - p_C)(1 - z)}{(p_C - p_E + (1 + p_C)z)(1 - p_C)z} = 1$$

$$(p_E - (1 - p_C)z)(1 - p_C)(1 - z) = (p_C - p_E + (1 + p_C)z)(1 - p_C)z$$

$$z = p_E$$

Checking the second derivatives of the terms in *H* shows that they are all negative because of the appearance of a $-(1 - p_C)^2$ factor, thereby confirming that *H* is at a maximum.

Substituting the value for *z* into the state probability solutions gives the following:

$$p_1 = p_E - (1 - p_C)p_E = p_C p_E$$

$$p_2 = p_C - p_E + (1 - p_C)\, p_E = p_C\,(1 - p_E)$$

$$p_3 = (1 - p_C)\, p_E$$

$$p_4 = (1 - p_C)(1 - p_E)$$

Substituting the value for *z* into the expression for the probability that C causes E gives the following:

$$P(C \Rightarrow E) = 1 - p_C + p_C p_E$$

The probability that C and E are coincident is equal to:

$$P(C \wedge E) = p_C \wedge p_E = p_1 = p_C p_E$$

$$\therefore M\&S\ Use\ Risk = P(C \Rightarrow E) \wedge P(C \wedge E) = (1 - p_C + p_C p_E) p_C p_E$$

Appendix 2: ASSIGNING PROBABILITY FOR MURM FACTORS

Each MURM factor or sub-factor is decomposed into a series of statements referred to as atoms. The truth of each of the statements and the set of possible states of the sub-factors defines the influence of each factor as interpreted in the field of information theory. This appendix presents the technique for evaluating the influence of each factor, the application of the technique to the taxonomy of factors, and the numerical results for each MURM factor. These results are used to consistently evaluate the probabilities associated with each factor.

A2.1 EVALUATING INFLUENCE USING INFORMATION THEORY

A means to assign the influence of state is required to be able to employ a state table in an overall analysis of utility. For the M&S use risk, the utility function is ultimately a probability relationship of realizing a risk (of inappropriate use and manifestation of undesirable consequence). There are several factors that contribute to this risk (viz., C1, C2, C3, and E), each with its own state table set, so the influences from each must ultimately be expressed in terms of a probability that being in a state will result in realizing the factor effect on the total question.

The issue then is how to assign a value for a factor's influence without introducing unintended bias. Approaches may range from subjective assignment to a precise mathematical relationship with or without justification. Data, if available, may determine these values (termed "weights" in many analyses, but termed "influences" in this discussion). In this MURM construct, knowledge is limited only to the number of states within a factor and the order of desirability (best to worst or worst to best) of them. A method from information theory provides a scheme to specify the relative influences of the states.

Forming Ordered States and Determining Their Influences Directly

The analysis of a state table is analogous to that of a cantilever beam, as illustrated in Figure A2-1. Consider the case of a factor that can only be in three states (A, B, and C) in order of most desirable to least desirable.

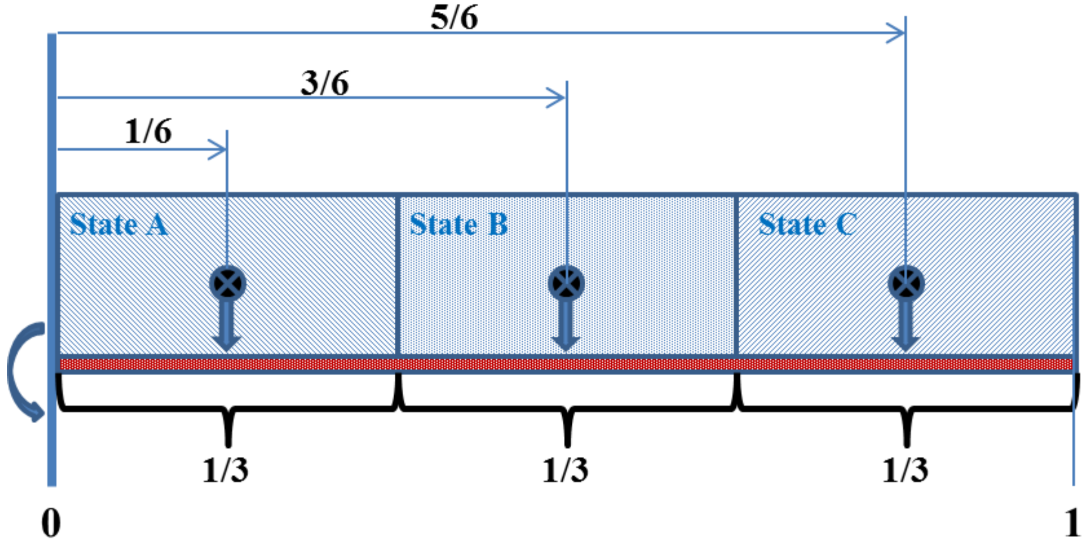

Figure A2-1: Three States with Preferential Order.

The three states must collectively span the entire length of the beam because they constitute the entire space allowed to the system. Suppose we consider each state a malleable mass of value; the influence of the state on the system may be represented by the physical moment of the weight of the mass on the supported (left-hand) side. This physical moment is equal to the weight of the mass multiplied by the distance to the centroid of interval spanned by the mass to the left side.

As drawn, State A's centroid is closest to the left side and, therefore, has the lowest moment; State B has the next lowest moment, and State C has the high moment. These physical moments are the relative influence of the state on the system. If the influence is interpreted as the probability that being in a state will result in realizing the factor, then the normalized location of the centroid (i.e., the beam has a total length of 1) corresponds to the probability of realizing the factor given the state, as drawn in Figure A2-1. Thus, we would assign P(Factor) = 1/6 if in State A, P(Factor) = ½ if in State B, or P(Factor) = 5/6 if in State C. The laydown of the masses on the beam are drawn such that each has a "footprint" equal to 1/3 the beam's length. Because we know only the number of states and their order of preference, an information theory argument forces us to consider each footprint equal or equal probability of distribution throughout the solution space. The information entropy is computed by the following equation:

$$S = \Sigma\, p_i \log_2 {}^{1}/_{p_i}$$

where p_i is the length of the footprint of the state on the beam. The maximum value of the information of the total system (i.e., the sum of the contribution of all states) is for this three-state illustration:

$$S_{max} = ⅓ \log_2 3 + ⅓ \log_2 3 + ⅓ \log_2 3 = 1.5849$$

The foregoing reasoning is extendible to n number of states, and the ratio of influences would be 1, 3, 5, …, (2n-1), and $S_{max} = \log_2(n)$. This counting technique (termed the "MIE count") is employed to determine the influences of other state configurations, especially in situations where more than one of the possible system states are considered equally preferable, as illustrated in Figure A2-2.

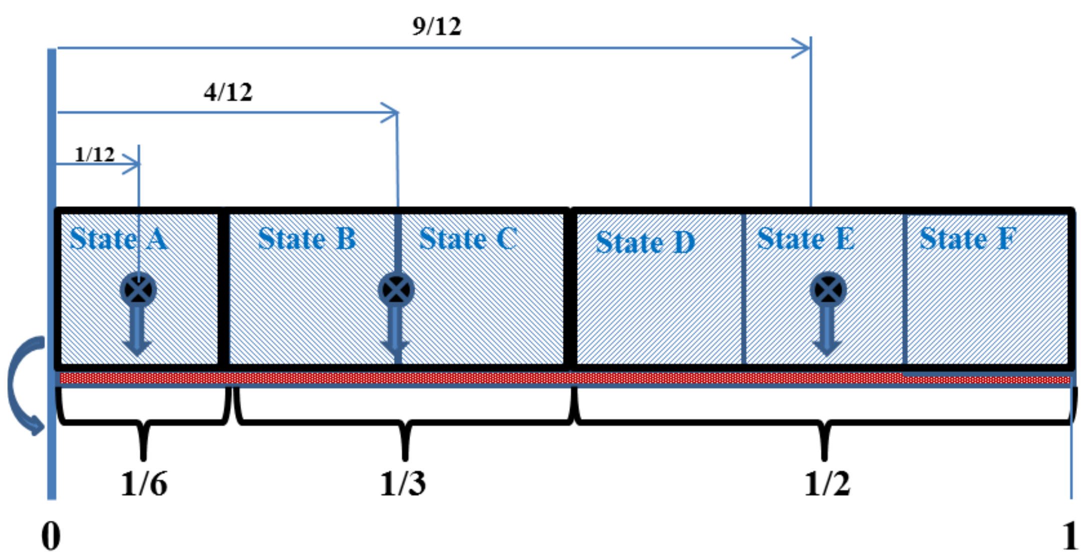

Figure A2-2: Six States with Three Preferential Order Groups.

In this system, there are six states (A, B, C, D, E, and F) arranged in a preferential order such that States B and C are considered equally preferable to each other and States D, E, and F are considered equally preferable to each other. This is equivalent to three states, where State A' is formed by State A, State B' is formed by States B and C, and State C' is formed by States D, E, and F. For those equally preferable states, their footprints are combined, and the combined centroid shifts to that corresponding to the median MIE count.

In Figure A2-2, the combination of the states from six separate states to three states results in changing the system information entropy from:

$$\log_2 6 = 2.545$$

To:

$$1/6 \log_2 6 + ⅓ \log_2 3 + ½ \log_2 2 = 1.4591$$

The decrease in total system entropy from 2.545 to 1.4591 is a measure of the information added to our knowledge of the system by adjusting our preferences of States B through F. Therefore, the influences of the factor's states are P(Factor) = 1/12 if in State A, P(Factor) = 1/3 if in State B or State C, and P(Factor) = ¾ if in State D, State E, or State F. This result is easily facilitated by the MIE count 1, (3, 5, median 4), (7, 9, 11, median 9) resulting in the ratio of influences as 1, 4, 9. If each is divided by the sum of the highest and lowest counts (1 + 11 = 12), the probability assignments for each grouping results: 1/12, 4/12, and 9/12.

Forming Ordered States and Determining Their Influences Indirectly

The approach of assigning P(Factor) applies to table constructs such as those for Factor C2 or Factor E. In those cases, the factor is embodied in only one overarching table. In cases where a factor is built from several sub-factors, the case of Factor C1 or Factor C3, a different approach is used to arrive at P(Factor). In these situations, the values of the influences contained in the sub-factors is carried to a utility function, and the cumulative distribution of the value of the utility function determines the assignment of P(Factor).

When combining the influences of several sub-factors, the form of the utility function may be indicated directly by the physics or mathematics of the system. However, as is often the case, the form of the utility function is *a prior* unavailable. In this circumstance, we use a maximum information entropy type of argument by imposing the condition that the utility function's form should be such that the influence of each sub-factor throughout the factor's solution space is independent of the other sub-factors. This leads to the functional equation of adding up the individual influences (after conditioning so that the internal ratios of the sub-factor influences are maintained, but the sums of each sub-factor's influences are equal). This is illustrated for the case of two sub-factors.

There are two factors, C_{x1} and C_{x2}, and the influence of each is represented by some function $F_{(Cx1)}$ and $F_{(Cx2)}$, respectively. The form of a utility function, U, that would meet the influence criterion is:

$$U = F(C_{x1}) + F(C_{x2})$$

because:

$$\frac{\partial U}{\partial F(C_{x1})} = \partial U / \partial F(C_{x1}) = 1$$

and:

$$\partial U / \partial F(C_{x2}) = 1$$

If a non-linear form such as $U = F(C_{x1}) \times F(C_{x2})$ were used instead, then $\partial U / \partial F(C_{x1}) = F(C_{x2})$ and $\partial U / \partial F(C_{x2}) = F(C_{x1})$. In this non-linear form, neither sub-factor C_{x1}'s influence nor sub-factor C_{x2}'s influence on U is dependent solely on its own function but is also dependent on the value of the other factor's influence.

Determine the Levels Associated with the Combinations of the Weights

Once the set of possible influence values have been defined using the approach outlined in the previous section:

1) Define all combinations of the T/F condition for each atom.
2) Evaluate the level associated with each combination.
3) Calculate the frequency for each level across all combinations.
4) Construct cumulative distribution from the frequencies.

A2.2 CALCULATING INFLUENCE FOR MURM FACTORS

The technique described has been applied to the MURM factors. It should be noted that this technique is independent of the factors themselves, and the M&S use risk equation depends only on the combined effect of the C-factors and the E-factors. Thus, if the sub-factors were to change, the influence and the utility function would need to be appropriately revised, but the M&S use risk equation and surface remain unchanged.

Factor C1: Clarity of Intended Use

For construction of a utility function leading to the assignment of P(C1), the values of influences for each sub-factor are adjusted so that the ratios between levels within each sub-factor are constant but the sum of the influences of each sub-factor is equal. This ensures that if the sub-factors do not have an equal number of levels, the sub-factors with a greater number of levels do not have a disproportionate influence over those with fewer. See Table A2-1.

Table A2-1: Table of Adjusted Influences for Factor C1 Sub-Factors.

Level	Sub-factor C_{11}	Sub-factor C_{12}	Sub-factor C_{13}	Sub-factor C_{14}
No. Levels	*2*	*3*	*4*	*4*
A	36	16	9	9
B	108	48	27	27
C		80	45	45
D			63	63
	144	*144*	*144*	*144*

For C1, the prior probabilities are derived from the cumulative distributions of utility function values formed by each respective set of sub-factors. The utility function is taken as the sum of the level influence multiplied by the number of levels within a factor for each of the factors. This ensures that the average contribution of each factor is comparable.

C1 Utility Function = SUM [No. Levels x Value Adjusted Influence] for C_{11} through C_{14}

The cumulative distribution of the C1 utility function values is shown in Figure A2-3.

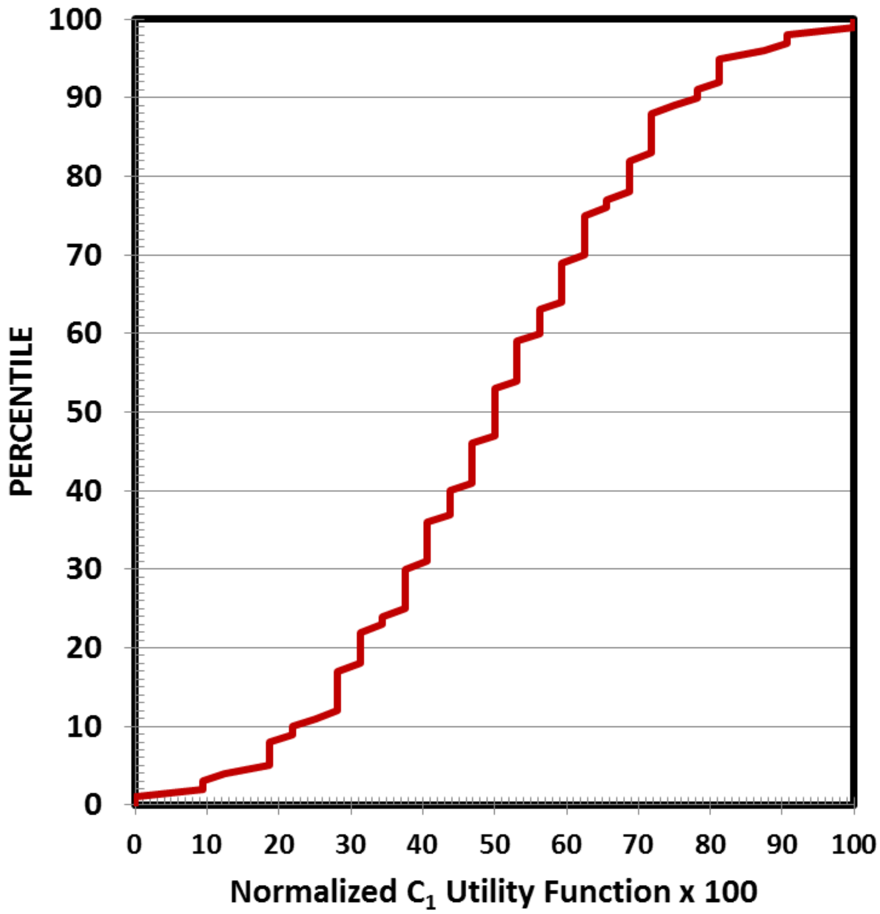

Figure A2-3: Cumulative Distribution of Normalized Utility Function for Factor C1.

P(C1) is ½ (Percentile/100) by assigning total ignorance (= 100 percentile), a probability of 0.5.

The maximum level of uncertainty (i.e., the maximum information entropy) of a binary decision of being correct is 50%. Therefore, for a completely unclear expression of intended use, the maximum probability that total ignorance will affect the inappropriate use is a 50 – 50 propositions. This assignment allows for a situation in which total ignorance of intended use neither adds to nor detracts from the probability of influencing inappropriate use of the M&S. Any departure (either up or down) from 0.5 implies additional knowledge that it will, or it will not.

Factor C2: M&S Limitations and Factor E: Effects

The construction of the Factor C2 state table and the Factor E state tables follow the foregoing procedure. Figure A2-4 depicts the diagram of the structure for the Factor C2 state table.

The 13 logical sentences contained in the C2 table, initially 13 separate states, are arranged into seven states (A through G) in Table A2-2. By grouping those sentences with equivalent influence, the entropy is reduced from its maximum value, $\log_2(13) = 3.700$ to a value of 2.661. The 1.039 bits decrease is a measure of the information obtained by sorting the 13 sentences from best to worst and grouping those with similar influence into a combined state.

Figure A2-5 depicts the diagram of the structure for the Factor E state table.

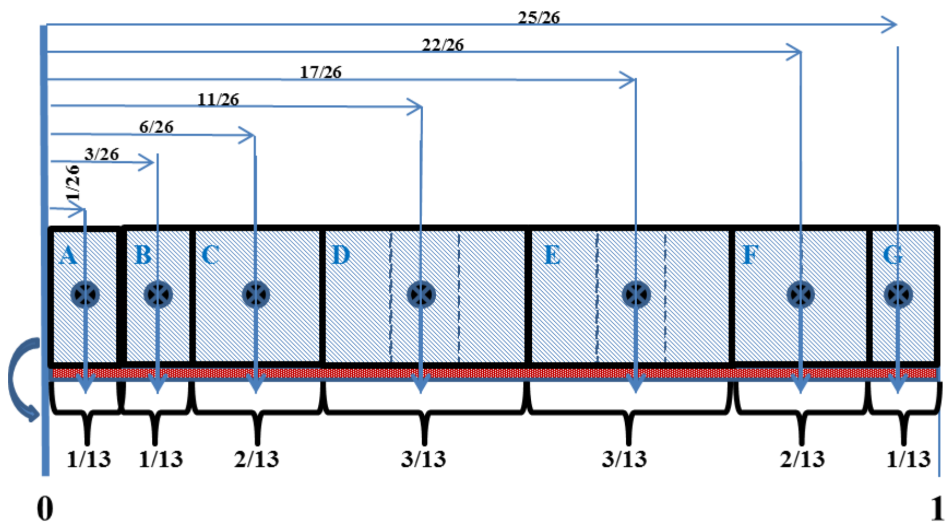

Figure A2-4: Diagram of the Factor C2 State Table.

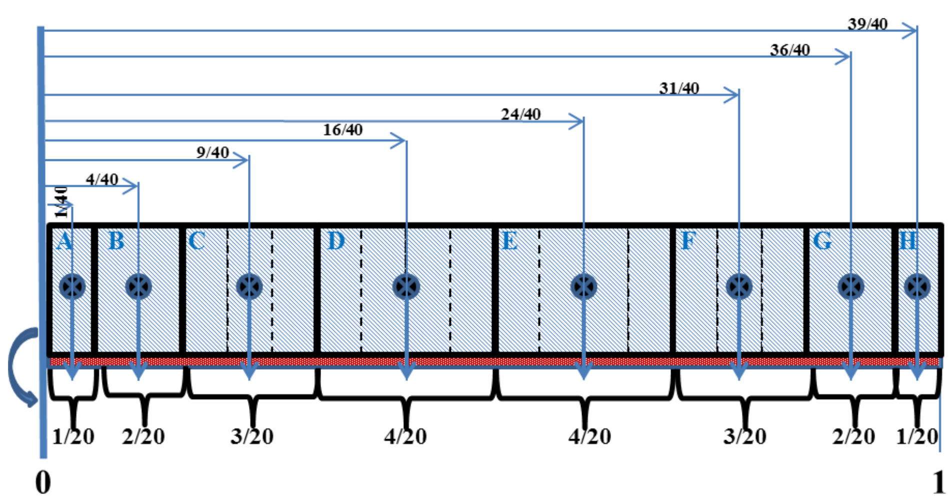

Figure A2-5: Diagram of the Factor E State Table.

Table A2-2: State Table for C2 M&S Limitations.

Factor Level	Consequence / Mitigation	Level Weighting	P(C2)
A	Negligible consequence / Mitigation not required	1	0.038
B	Negligible consequence / Mitigation complete	3	0.115
C	Negligible consequence / Mitigation partial or Minor consequence / Mitigation complete	6	0.231
D	Negligible consequence / Mitigation impossible or Minor consequence / Mitigation partial or Serious consequence / Mitigation complete	11	0.423
E	Minor consequence / Mitigation impossible or Serious consequence / Mitigation partial or Grave consequence / Mitigation complete	17	0.654
F	Serious consequence / Mitigation impossible or Grave consequence / Mitigation partial	22	0.846
G	Grave consequence / Mitigation impossible	25	0.962

The 20 logical sentences contained in the Factor E table are arranged into eight states (A through H), reducing the partition entropy from its maximum value, $\log_2(20) = 4.322$, to a value of 2.846. The 1.475 bits decrease is a measure of the information obtained by the preference sorting of the 20 sentences, initially separate states, from best to worst and grouping those with similar influence into the reduced number of states, some of which contain more than one sentence.

Table A2-3 is reformed into Table A2-4, from which the MIE count on this new structure is performed and the P(E) is designated for the eight states into which the 20 combinations fall.

Table A2-3: Value of Utility of the Two Factors Within the Full Space.

<table>
<tr><td colspan="4"></td><td colspan="4">M&S Reliance</td></tr>
<tr><td colspan="4"></td><td>A</td><td>B</td><td>C</td><td>D</td></tr>
<tr><td colspan="3"></td><td>Reliance
Influence →</td><td>1</td><td>3</td><td>5</td><td>7</td></tr>
<tr><td colspan="2"></td><td>Impact
Influence
↓</td><td>Description</td><td>Supplemental</td><td>Secondary</td><td>Primary</td><td>Only</td></tr>
<tr><td rowspan="5">M&S
Impact
(Rows)</td><td>A</td><td>1</td><td>Single Low Risk Area</td><td>2</td><td>4</td><td>6</td><td>8</td></tr>
<tr><td>B</td><td>3</td><td>Single Medium Risk Area</td><td>4</td><td>6</td><td>8</td><td>10</td></tr>
<tr><td>C</td><td>5</td><td>Multiple Medium or Low Risk Areas</td><td>6</td><td>8</td><td>10</td><td>12</td></tr>
<tr><td>D</td><td>7</td><td>Single High Risk Area</td><td>8</td><td>10</td><td>12</td><td>14</td></tr>
<tr><td>E</td><td>9</td><td>Multiple High Risk Areas</td><td>10</td><td>12</td><td>14</td><td>16</td></tr>
</table>

Table A2-4: Complete Analysis of Influence of All Allowable Combinations of Factor E.

<table>
<tr><th>Level</th><th>Sum of Influences</th><th>Column-Row (Table A2-3)</th><th colspan="2">Descriptions
Reliance / Impact Area(s)</th><th>MIE Count</th><th>P(E)</th></tr>
<tr><td>A</td><td>2</td><td>A-A</td><td>Supplemental</td><td>Single Low Risk</td><td>1</td><td>0.025</td></tr>
<tr><td>B</td><td>4</td><td>A-B
B-A</td><td>Supplemental
Secondary</td><td>Single Medium Risk
Single Low Risk</td><td>4</td><td>0.100</td></tr>
<tr><td>C</td><td>6</td><td>A-C
B-B
C-A</td><td>Supplemental
Secondary
Primary</td><td>Multiple Med / Low Risks
Single Medium Risk
Single Low Risk</td><td>9</td><td>0.225</td></tr>
<tr><td>D</td><td>8</td><td>A-D
B-C
C-B
D-A</td><td>Supplemental
Secondary
Primary
Only</td><td>Single High Risk
Multiple Med / Low Risk
Single Medium Risk
Single Low Risk</td><td>16</td><td>0.400</td></tr>
</table>

Level	Sum of Influences	Column-Row (Table A2-3)	Descriptions Reliance / Impact Area(s)		MIE Count	P(E)
E	10	A-E	Supplemental	Multiple High Risk	24	0.600
		B-D	Secondary	Single High Risk		
		C-C	Primary	Multiple Med / Low Risk		
		D-B	Only	Single Medium Risk		
F	12	B-E	Secondary	Multiple High Risk	31	0.775
		C-D	Primary	Single High Risk		
		D-C	Only	Multiple Med / Low Risk		
G	14	C-E	Primary	Multiple High Risk	36	0.900
		D-D	Only	Single High Risk		
H	16	D-E	Only	Multiple High Risk	39	0.975

Factor C3: Verification and Validation Confidence

Table A2-5: Verification and Validation Confidence: State Tables of Sub-Factors.

Sub-Factor C_{31}: Input Validation Analysis		
Level	**Sentence**	**Influence**
A	$(u \wedge v \wedge w \wedge x \wedge y \wedge z)$	1
B	(5 of 6) and no objective-threshold conflicts	5
C	(4 of 6) and no objective-threshold conflicts	14
D	(3 of 6) and no objective-threshold conflicts	27
E	(2 of 6) and no objective-threshold conflicts	40
F	(1 of 6) and no objective-threshold conflicts	49
G	(0 of 6)	53
Sub-Factor C_{32}: Verification Analysis		
Level	**Sentence**	**Influence**
A	$(u \wedge v \wedge w \wedge x \wedge y \wedge z)$	1
B	(5 of 6) and no objective-threshold conflicts	5
C	(4 of 6) and no objective-threshold conflicts	14
D	(3 of 6) and no objective-threshold conflicts	27
E	(2 of 6) and no objective-threshold conflicts	40
F	(1 of 6) and no objective-threshold conflicts	49
G	(0 of 6)	53

Sub-Factor C_{33}: Validation Analysis		
Level	**Sentence**	**Influence**
A	(u∧v∧w∧x∧y∧z)	1
B	(5 of 6) and no objective-threshold conflicts	5
C	(4 of 6) and no objective-threshold conflicts	14
D	(3 of 6) and no objective-threshold conflicts	27
E	(2 of 6) and no objective-threshold conflicts	40
F	(1 of 6) and no objective-threshold conflicts	49
G	(0 of 6)	53
Key		
u ≡ Scope meets objective v ≡ Scope meets threshold w ≡ Referent meets objective (C31 , C33); Framework meets objective (C32) x ≡ Referent meets threshold (C31 , C33); Framework meets threshold (C32) y ≡ Methodology meets objective z ≡ Methodology meets threshold		

Table A2-6: Table of Adjusted Influences for Factor C3 Sub-Factors.

Level	Sub-Factor C31	Sub-Factor C32	Sub-Factor C33
No. Levels	7	7	7
A	1	1	1
B	5	5	5
C	14	14	14
D	27	27	27
E	40	40	40
F	49	49	49
G	53	53	53
	189	189	189

C3 Utility Function = SUM [No. Levels x Value Adjusted Influence] for C_{31} through C_{33}.

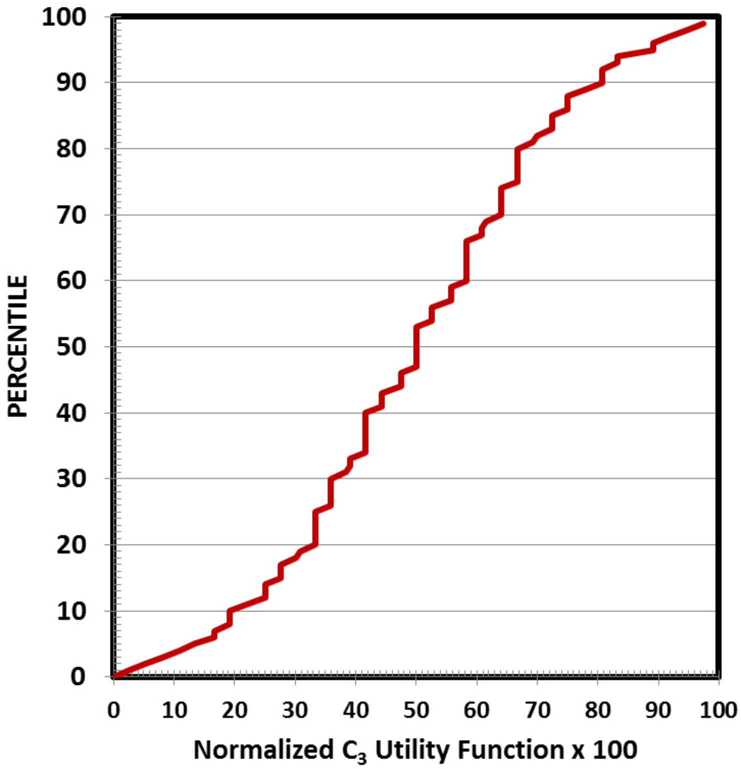

Figure A2-6: Cumulative Distribution of Normalized Utility Function for Factor C3.

P(C3) is ½ (Percentile/100) by assigning total ignorance (= 100 percentile) a probability of 0.5.

The maximum level of uncertainty (i.e., the maximum information entropy) of a binary decision of being correct is 50%. Therefore, for a completely unclear expression of intended use, the maximum probability that total ignorance will affect the inappropriate use is a 50-50 proposition. This assignment allows for a situation in which total ignorance of intended use neither adds to nor detracts from the probability of influencing inappropriate use of the M&S. Any departure (either up or down) from 0.5 implies additional knowledge that it will, or it will not.

Factor E: Effects

Factor E is derived in a manner similar to Factor C2.

Appendix 3: EXPLANATION OF ISOCLINE BEHAVIOR

Figure A3-1 shows the two-dimensional projection of the M&S use risk (i.e., the M&S use risk space) isoclines (lines of equal M&S use risk), as discussed in Section 3.3. As noted previously, the graph is visually counterintuitive. Specifically, in the lower right quadrant (high values of p_C and low values of p_E), the M&S use risk increases.

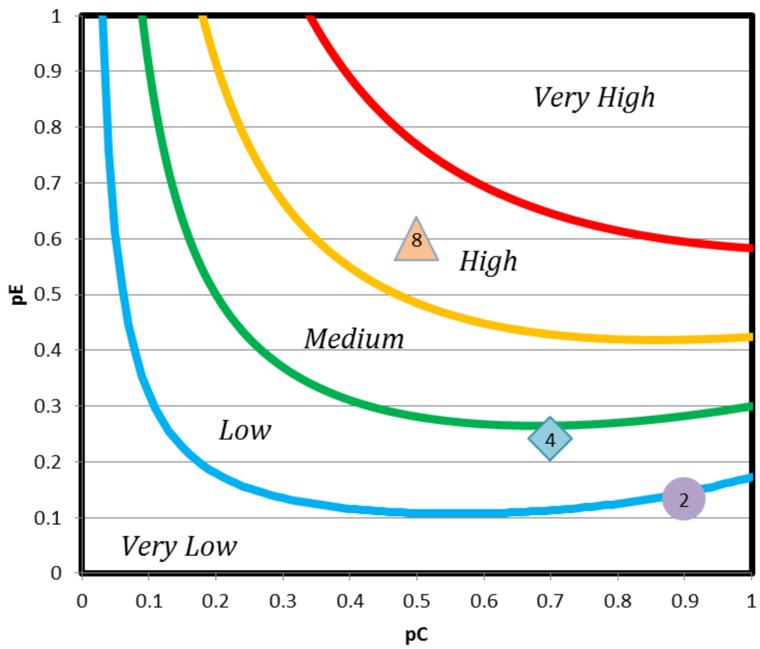

Figure A3-1: Two-Dimensional Projection of the Risk Constellation.

The explanation of this phenomenon can be drawn directly from the definition of entropy in information theory. In information theory, the binary entropy function $H(p)$ is defined as the entropy of a binary event with probability p, as follows:

$$H(p) = -p \cdot \log_2(p) - (1-p) \cdot \log_2(1-p)$$

The graph for this function is depicted in Figure A3-2.

The M&S use risk is a probability that is compounded from several, more elemental, probabilities. As such, the M&S use risk is a binary event. This means the equation and its graph (Figure A3-2) are still valid. The graph shows that the entropy is maximum (least amount of information / least amount of certainty) when the probability is 0.5. The entropy diminishes on either side of this maximum, and thus values farther away from 0.5 provide more certainty regarding the probability.

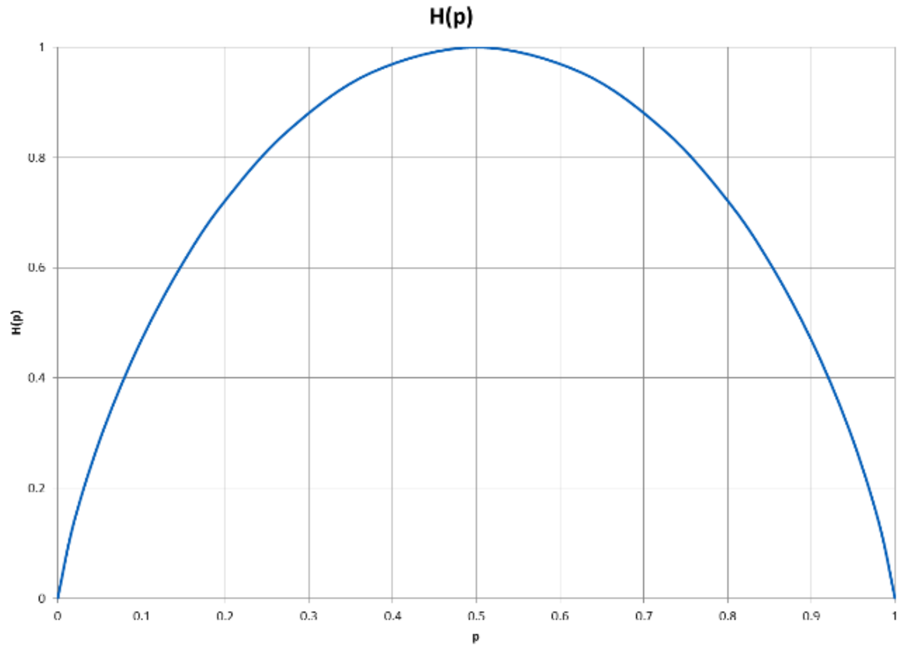

Figure A3-2: M&S Use Risk Entropy.

The M&S use risk space is defined in terms of p_C and p_E. This is also true for the information entropy. In the case of M&S use risk, the information entropy equation can be rewritten as a function of p_C and p_E, as follows:

$$H(p) = -p \cdot \log_2(p) - (1-p) \cdot \log_2(1-p)$$

$$H(UR) = -UR \cdot \log_2(UR) - (1-UR) \cdot \log_2(1-UR)$$

$$H(UR) = -P(C \wedge E) \cdot P(C \Rightarrow \text{E}) \cdot \log_2\big(P(C \wedge E) \cdot P(C \Rightarrow \text{E})\big) - \big(1 - P(C \wedge E) \cdot P(C \Rightarrow \text{E})\big) \cdot \log_2\big(1 - P(C \wedge E) \cdot P(C \Rightarrow \text{E})\big)$$

$$H(UR) = -p_C \cdot p_E \cdot [1 - p_E + p_C \cdot p_E] \cdot \log_2[p_C \cdot p_E \cdot (1 - p_E + p_C \cdot p_E)] - [1 - p_C \cdot p_E \cdot (1 - p_E + p_C \cdot p_E)] \cdot \log_2[1 - p_C \cdot p_E \cdot (1 - p_E + p_C \cdot p_E)]$$

This, albeit a very long and cumbersome equation, can easily be evaluated as a function of both p_C and p_E to determine the conditions that provide the maximum entropy. The proper way to do this would involve taking second derivatives of the above expression and solving some complicated algebraic equations. This is tedious to say the least and not entirely necessary for the purpose of exploring the behavior of the M&S use risk function. Rather, the Excel solver features to find both global and local information entropy maxima within the ranges for both p_C and p_E fixing one of the parameters at a time was used.

The curve for maximum information entropy (labelled MAX[H(UR)]) is shown with the M&S use risk isoclines in Figure A3-3. This curve represents a local maximum for any constant value of p_E. The global maximum on this curve in the top-right quadrant of Figure A3-3 coincides with $UR = 0.5$, $H(UR) = 1$. This global maximum is expected, as predicted by the information entropy function; see Figure A3-2.

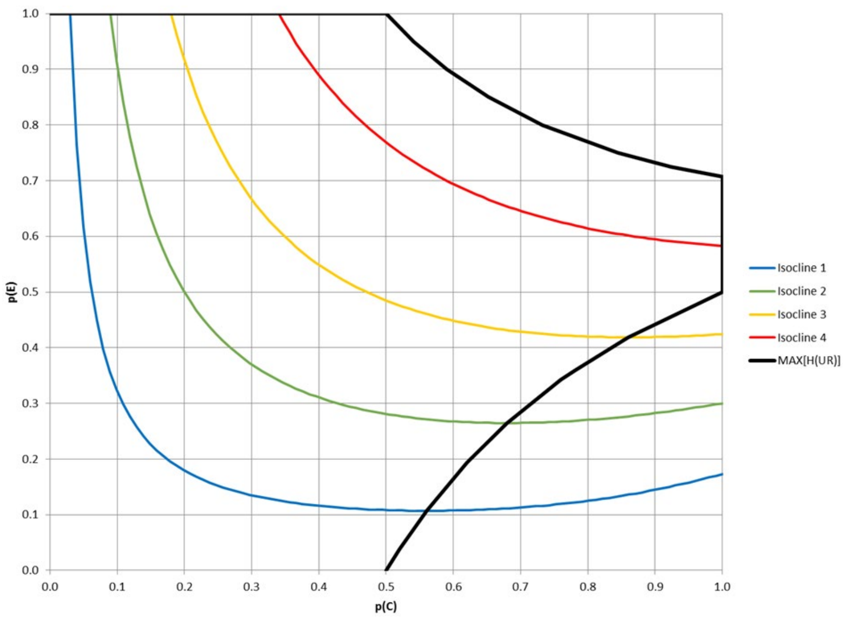

Figure A3-3: Effect of Variations in Entropy.

The most confusing case occurs for a constant low p_E. A trajectory beginning on point 1 in Figure A3-4 (highest p_C) and progressing down p_C crosses the green isocline (Isocline 3) on 4 twice. In terms of UR, this translates in a brief increase (~14%), followed by a continuous decrease in M&S use risk.

Table A3-1 and Figure A3-4 show the brief increase in UR. In terms of information entropy, point 1 is in a low entropy state on the right side of a local maximum. As p_C decreases (moving toward the left on Figure A3-4), the entropy approaches a local maximum for this constant value of p_E. Point 2 is the condition of maximum entropy. An increase in information entropy is understood as a reduction in certainty regarding the probability and in this case is presented as an increase in UR. After the point of maximum entropy, the entropy reduces approaching point 3. As the information entropy reduces, the certainty is increased and as a consequence UR is reduced.

Finally, recall that the M&S use risk is calculated as a combination of probabilities. The maximum information entropy for M&S use risk does not necessarily occur for the maximum entropy of any of the probabilities considered. However, in this case, the local maximum in the lower right of Figure A3-4 coincides with $P(C \Rightarrow \mathrm{E}) = 0.5$, $H\big(P(C \Rightarrow \mathrm{E})\big) = 1$. Moving from point 1 to point 2 represents losing information regarding the relationship between the two probabilities; that is, less can be inferred about the C-factors causing the E-factors. As a consequence of this loss of information, the M&S use risk is increased.

Table A3-1: Comparison Points.

Point #	pC	pE	M&S Use Risk	H(M&S Use Risk)
1	0.941	0.280	0.085	0.420
2	0.694	0.280	0.097	0.460
3	0.448	0.280	0.085	0.420

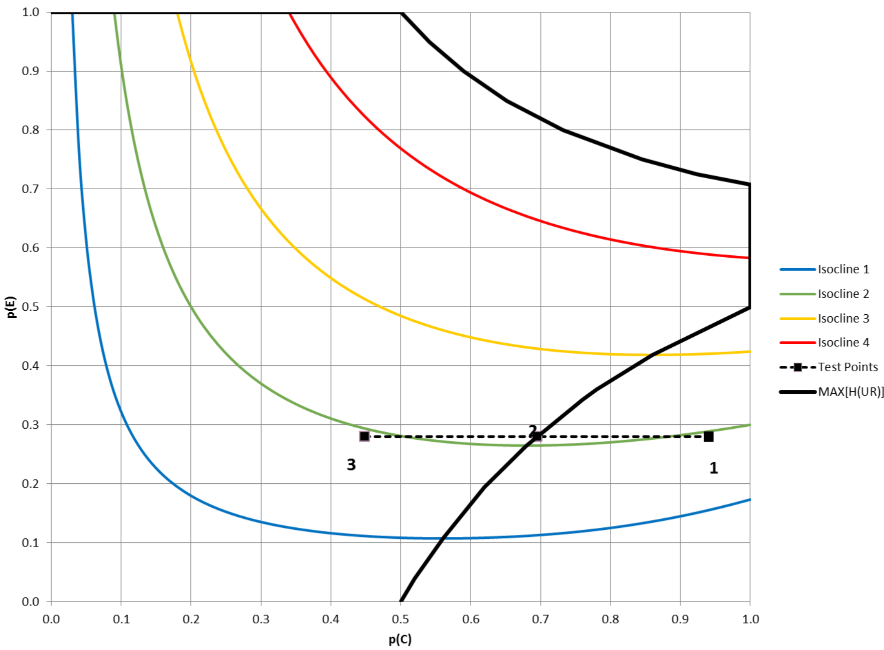

Figure A3-4: Changes in M&S Use Risk for Constant P(E).

REPORT DOCUMENTATION PAGE

1. Recipient's Reference	2. Originator's References	3. Further Reference	4. Security Classification of Document
	STO-TR-MSG-139 AC/323(MSG-139)TP/945	ISBN 978-92-837-2269-4	PUBLIC RELEASE

5. Originator

Science and Technology Organization
North Atlantic Treaty Organization
BP 25, F-92201 Neuilly-sur-Seine Cedex, France

6. Title

Guidelines for Modelling and Simulation (M&S) Use Risk Identification, Analysis, and Mitigation

7. Presented at/Sponsored by

Final Report of Modelling and Simulation Group MSG-139.

8. Author(s)/Editor(s)	9. Date
Multiple	September 2021

10. Author's/Editor's Address	11. Pages
Multiple	90

12. Distribution Statement

There are no restrictions on the distribution of this document. Information about the availability of this and other STO unclassified publications is given on the back cover.

13. Keywords/Descriptors

Modeling; Simulation; Use risk; Use risk analysis; Use risk identification; Use risk mitigation; Validation; Verification

14. Abstract

M&S use risk is recognized as a critical component to M&S assessment and use. Despite the consensus in the M&S community on the importance of this topic, there are no accepted methods available for qualification or quantification of M&S use risk that account for project-specific M&S requirements and constraints. The primary objective of NATO Modelling and Simulation Group (NMSG) 139 was the assessment of methodologies that would support Modelling and Simulation (M&S) use risk identification, management, and mitigation. Effective management of risk requires both identification of the risk and a means by which to balance investments to mitigate them. Such an evaluation is made based on an assessment of the likelihood of the realization of the risk and the impact of that realization. When the risks are identified and assessed, mitigation strategies can be developed. The M&S Use Risk Methodology (MURM) was identified early on as a framework that addressed the identification, management, and mitigation of risk. This document reports the results of MSG-139 efforts with a specific focus on the formalization and application of the MURM.

NORTH ATLANTIC TREATY ORGANIZATION

NATO OTAN

BP 25

F-92201 NEUILLY-SUR-SEINE CEDEX • FRANCE

Télécopie 0(1)55.61.22.99 • E-mail mailbox@cso.nato.int

SCIENCE AND TECHNOLOGY ORGANIZATION

DIFFUSION DES PUBLICATIONS STO NON CLASSIFIEES

Les publications de l'AGARD, de la RTO et de la STO peuvent parfois être obtenues auprès des centres nationaux de distribution indiqués ci-dessous. Si vous souhaitez recevoir toutes les publications de la STO, ou simplement celles qui concernent certains Panels, vous pouvez demander d'être inclus soit à titre personnel, soit au nom de votre organisation, sur la liste d'envoi.

Les publications de la STO, de la RTO et de l'AGARD sont également en vente auprès des agences de vente indiquées ci-dessous.

Les demandes de documents STO, RTO ou AGARD doivent comporter la dénomination « STO », « RTO » ou « AGARD » selon le cas, suivi du numéro de série. Des informations analogues, telles que le titre est la date de publication sont souhaitables.

Si vous souhaitez recevoir une notification électronique de la disponibilité des rapports de la STO au fur et à mesure de leur publication, vous pouvez consulter notre site Web (http://www.sto.nato.int/) et vous abonner à ce service.

CENTRES DE DIFFUSION NATIONAUX

ALLEMAGNE
Streitkräfteamt / Abteilung III
Fachinformationszentrum der Bundeswehr (FIZBw)
Gorch-Fock-Straße 7, D-53229 Bonn

BELGIQUE
Royal High Institute for Defence – KHID/IRSD/RHID
Management of Scientific & Technological Research for Defence, National STO Coordinator
Royal Military Academy – Campus Renaissance
Renaissancelaan 30, 1000 Bruxelles

BULGARIE
Ministry of Defence
Defence Institute "Prof. Tsvetan Lazarov"
"Tsvetan Lazarov" bul no.2
1592 Sofia

CANADA
DGSIST 2
Recherche et développement pour la défense Canada
60 Moodie Drive (7N-1-F20)
Ottawa, Ontario K1A 0K2

DANEMARK
Danish Acquisition and Logistics Organization (DALO)
Lautrupbjerg 1-5
2750 Ballerup

ESPAGNE
Área de Cooperación Internacional en I+D
SDGPLATIN (DGAM)
C/ Arturo Soria 289
28033 Madrid

ESTONIE
Estonian National Defence College
Centre for Applied Research
Riia str 12
Tartu 51013

ETATS-UNIS
Defense Technical Information Center
8725 John J. Kingman Road
Fort Belvoir, VA 22060-6218

FRANCE
O.N.E.R.A. (ISP)
29, Avenue de la Division Leclerc
BP 72
92322 Châtillon Cedex

GRECE (Correspondant)
Defence Industry & Research General Directorate, Research Directorate
Fakinos Base Camp, S.T.G. 1020
Holargos, Athens

HONGRIE
Hungarian Ministry of Defence
Development and Logistics Agency
P.O.B. 25
H-1885 Budapest

ITALIE
Ten Col Renato NARO
Capo servizio Gestione della Conoscenza
F. Baracca Military Airport "Comparto A"
Via di Centocelle, 301
00175, Rome

LUXEMBOURG
Voir Belgique

NORVEGE
Norwegian Defence Research Establishment
Attn: Biblioteket
P.O. Box 25
NO-2007 Kjeller

PAYS-BAS
Royal Netherlands Military Academy Library
P.O. Box 90.002
4800 PA Breda

POLOGNE
Centralna Biblioteka Wojskowa
ul. Ostrobramska 109
04-041 Warszawa

PORTUGAL
Estado Maior da Força Aérea
SDFA – Centro de Documentação
Alfragide
P-2720 Amadora

REPUBLIQUE TCHEQUE
Vojenský technický ústav s.p.
CZ Distribution Information Centre
Mladoboleslavská 944
PO Box 18
197 06 Praha 9

ROUMANIE
Romanian National Distribution Centre
Armaments Department
9-11, Drumul Taberei Street
Sector 6
061353 Bucharest

ROYAUME-UNI
Dstl Records Centre
Rm G02, ISAT F, Building 5
Dstl Porton Down
Salisbury SP4 0JQ

SLOVAQUIE
Akadémia ozbrojených síl gen. M.R. Štefánika, Distribučné a informačné stredisko STO
Demänová 393
031 01 Liptovský Mikuláš 1

SLOVENIE
Ministry of Defence
Central Registry for EU & NATO
Vojkova 55
1000 Ljubljana

TURQUIE
Milli Savunma Bakanlığı (MSB)
ARGE ve Teknoloji Dairesi Başkanlığı
06650 Bakanliklar – Ankara

AGENCES DE VENTE

The British Library Document Supply Centre
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
ROYAUME-UNI

Canada Institute for Scientific and Technical Information (CISTI)
National Research Council Acquisitions
Montreal Road, Building M-55
Ottawa, Ontario K1A 0S2
CANADA

Les demandes de documents STO, RTO ou AGARD doivent comporter la dénomination « STO », « RTO » ou « AGARD » selon le cas, suivie du numéro de série (par exemple AGARD-AG-315). Des informations analogues, telles que le titre et la date de publication sont souhaitables. Des références bibliographiques complètes ainsi que des résumés des publications STO, RTO et AGARD figurent dans le « NTIS Publications Database » (http://www.ntis.gov).

NORTH ATLANTIC TREATY ORGANIZATION

BP 25
F-92201 NEUILLY-SUR-SEINE CEDEX • FRANCE
Télécopie 0(1)55.61.22.99 • E-mail mailbox@cso.nato.int

SCIENCE AND TECHNOLOGY ORGANIZATION

DISTRIBUTION OF UNCLASSIFIED STO PUBLICATIONS

AGARD, RTO & STO publications are sometimes available from the National Distribution Centres listed below. If you wish to receive all STO reports, or just those relating to one or more specific STO Panels, they may be willing to include you (or your Organisation) in their distribution.

STO, RTO and AGARD reports may also be purchased from the Sales Agencies listed below.

Requests for STO, RTO or AGARD documents should include the word 'STO', 'RTO' or 'AGARD', as appropriate, followed by the serial number. Collateral information such as title and publication date is desirable.

If you wish to receive electronic notification of STO reports as they are published, please visit our website (http://www.sto.nato.int/) from where you can register for this service.

NATIONAL DISTRIBUTION CENTRES

BELGIUM
Royal High Institute for Defence – KHID/IRSD/RHID
Management of Scientific & Technological Research for Defence, National STO Coordinator
Royal Military Academy – Campus Renaissance
Renaissancelaan 30
1000 Brussels

BULGARIA
Ministry of Defence
Defence Institute "Prof. Tsvetan Lazarov"
"Tsvetan Lazarov" bul no.2
1592 Sofia

CANADA
DSTKIM 2
Defence Research and Development Canada
60 Moodie Drive (7N-1-F20)
Ottawa, Ontario K1A 0K2

CZECH REPUBLIC
Vojenský technický ústav s.p.
CZ Distribution Information Centre
Mladoboleslavská 944
PO Box 18
197 06 Praha 9

DENMARK
Danish Acquisition and Logistics Organization (DALO)
Lautrupbjerg 1-5
2750 Ballerup

ESTONIA
Estonian National Defence College
Centre for Applied Research
Riia str 12
Tartu 51013

FRANCE
O.N.E.R.A. (ISP)
29, Avenue de la Division Leclerc – BP 72
92322 Châtillon Cedex

GERMANY
Streitkräfteamt / Abteilung III
Fachinformationszentrum der Bundeswehr (FIZBw)
Gorch-Fock-Straße 7
D-53229 Bonn

GREECE (Point of Contact)
Defence Industry & Research General Directorate, Research Directorate
Fakinos Base Camp, S.T.G. 1020
Holargos, Athens

HUNGARY
Hungarian Ministry of Defence
Development and Logistics Agency
P.O.B. 25
H-1885 Budapest

ITALY
Ten Col Renato NARO
Capo servizio Gestione della Conoscenza
F. Baracca Military Airport "Comparto A"
Via di Centocelle, 301
00175, Rome

LUXEMBOURG
See Belgium

NETHERLANDS
Royal Netherlands Military Academy Library
P.O. Box 90.002
4800 PA Breda

NORWAY
Norwegian Defence Research Establishment, Attn: Biblioteket
P.O. Box 25
NO-2007 Kjeller

POLAND
Centralna Biblioteka Wojskowa
ul. Ostrobramska 109
04-041 Warszawa

PORTUGAL
Estado Maior da Força Aérea
SDFA – Centro de Documentação
Alfragide
P-2720 Amadora

ROMANIA
Romanian National Distribution Centre
Armaments Department
9-11, Drumul Taberei Street
Sector 6
061353 Bucharest

SLOVAKIA
Akadémia ozbrojených síl gen M.R. Štefánika, Distribučné a informačné stredisko STO
Demänová 393
031 01 Liptovský Mikuláš 1

SLOVENIA
Ministry of Defence
Central Registry for EU & NATO
Vojkova 55
1000 Ljubljana

SPAIN
Área de Cooperación Internacional en I+D
SDGPLATIN (DGAM)
C/ Arturo Soria 289
28033 Madrid

TURKEY
Milli Savunma Bakanlığı (MSB)
ARGE ve Teknoloji Dairesi Başkanlığı
06650 Bakanliklar – Ankara

UNITED KINGDOM
Dstl Records Centre
Rm G02, ISAT F, Building 5
Dstl Porton Down, Salisbury SP4 0JQ

UNITED STATES
Defense Technical Information Center
8725 John J. Kingman Road
Fort Belvoir, VA 22060-6218

SALES AGENCIES

The British Library Document Supply Centre
Boston Spa, Wetherby
West Yorkshire LS23 7BQ
UNITED KINGDOM

Canada Institute for Scientific and Technical Information (CISTI)
National Research Council Acquisitions
Montreal Road, Building M-55
Ottawa, Ontario K1A 0S2
CANADA

Requests for STO, RTO or AGARD documents should include the word 'STO', 'RTO' or 'AGARD', as appropriate, followed by the serial number (for example AGARD-AG-315). Collateral information such as title and publication date is desirable. Full bibliographical references and abstracts of STO, RTO and AGARD publications are given in "NTIS Publications Database" (http://www.ntis.gov).

ISBN 978-92-837-2269-4