ACCELERATING THE BROAD IMPLEMENTATION OF

VERIFICATION & VALIDATION

IN COMPUTATIONAL MODELS OF THE MECHANICS OF MATERIALS AND STRUCTURES

TMS A Study Organized by The Minerals, Metals & Materials Society

Accelerating the Broad Implementation of Verification & Validation in Computational Models of the Mechanics of Materials and Structures

ACCELERATING THE BROAD IMPLEMENTATION OF

VERIFICATION & VALIDATION

IN COMPUTATIONAL MODELS OF THE MECHANICS OF MATERIALS AND STRUCTURES

A STUDY ORGANIZED BY

The Minerals, Metals & Materials Society (TMS)

www.tms.org/VVaccelerator

Design: Cover and layout design by Bob Demmler, TMS.
Three dimensional microstructure cover image (bottom, left) provided courtesy of Dr. David Rowenhorst, Naval Research Laboratory.

Electronic copies of this report are available online at www.tms.org/VVaccelerator.

DOI: 10.7449/VandV_2 ISBN: 978-0-578-75450-5

The Minerals, Metals & Materials Society (TMS)

Promoting the global science and engineering professions concerned with minerals, metals, and materials

The Minerals, Metals & Materials Society (TMS) is a member-driven, international organization dedicated to the science and engineering professions concerned with minerals, metals and materials. TMS includes approximately 14,000 professional and student members from more than 80 countries, representing industry, government and academia.

The society's technical focus spans a broad range—from minerals processing and primary metals production to basic research and the advanced applications of materials.

In recent years, TMS has established itself as a leader in advancing integrated computational materials engineering (ICME), computational materials science and engineering, multiscale materials modeling and simulation, materials data infrastructure issues, and advanced manufacturing methodologies.

To facilitate global knowledge exchange and networking, TMS organizes meetings; develops continuing education courses; publishes conference proceedings, peer-reviewed journals, and textbooks; develops science and technology accelerator studies and reports; and presents a variety of web resources, accessed through www.tms.org.

TMS also represents materials science and engineering professions in the accreditation of educational programs and in the registration of professional engineers across the United States.

A recognized leader in bridging the gap between materials research and application, TMS leads and enables advancements in a broad spectrum of domestic and global initiatives.

TMS is committed to advancing diversity in the minerals, metals, and materials professions and to promoting an inclusive professional culture that welcomes and engages all who seek to contribute to the field.

www.tms.org

TMS

The Minerals, Metals & Materials Society

Contents

Acknowledgements xi

Preface xxi

Executive Summary xxv

I. Background, Motivation, and Study Process 1

II. Value Proposition of V&V 5

III. Recommended Practices for V&V and UQ of Computational Models Associated with the Mechanics of Materials and Structures 13

III.1. Framework, Definitions, and Overview 15

III.2. Recommended Practices – Process Steps, Tasks, and Contributors 22

IV. Challenge Problems and Sustainable Funding Programs 37

V. Symposia, Conferences, Workshops, and Studies 47

V.1. Symposia and/or Conferences 47

V.2. Studies/Workshops 49

VI. Strategies to Connect Modelers and Experimentalists 53

VII. Strategies for Improving Multidisciplinary Training and Curricula Development 59

VIII. Final Comments and Call to Action 65

References 69

Appendix: Glossary 77

Acknowledgements

This report on *Accelerating the Broad Implementation of Verification and Validation in Computational Models of the Mechanics of Materials and Structures* is a culmination of the efforts of a group of internationally recognized subject matter experts from academia, industry, and government, who volunteered significant time to this study. Collectively, these experts participated in several online meetings and teleconferences, attended in-person and/or virtual facilitated workshops, and helped to write, edit, and review the report.

Their leadership, dedication, and involvement were foundational to this effort. We want to express our sincere gratitude for their hard work and contributions that are sure to have a lasting impact on the community. We trust that others will follow their lead and, after reading this document, identify how they, too, can propel the field forward through the adoption, development, and implementation of verification and validation practices within their workflows and fields of practice.

William L. Oberkampf, Study Team Chair
George Spanos, Project Leader

Study Team Members

William L. Oberkampf (Study Team Chair)
Adjunct Faculty in the Crofton Department of Aerospace and Ocean Engineering
Virginia Polytechnic Institute and State University

William Oberkampf is an adjunct faculty member in the Kevin T. Crofton Department of Aerospace and Ocean Engineering at Virginia Tech, Blacksburg, Virginia. He is also an engineering consultant working with various government laboratories and business organizations in both the U.S. and Europe. Oberkampf specializes in issues concerning computational modeling and simulation, particularly verification, model validation and experimental activities, model parameter calibration, sensitivity analyses, and uncertainty quantification. He has taught over 50 short courses on these topics in the U.S., Europe, and China, and has published over 180 journal articles, book chapters, and conference papers. Oberkampf is the co-author of *Verification and Validation in Scientific Computing*, Cambridge University Press, 2010. He served in staff and management positions at Sandia National Laboratories over a period of 29 years. Early in his career, Oberkampf was a faculty member in the Department of Mechanical Engineering at the University of Texas at Austin. He earned his Ph.D. in aerospace engineering from the University of Notre Dame in 1970, receiving the Outstanding Engineering Graduate Award. Oberkampf is a Fellow of the American Institute of Aeronautics and Astronautics and Technical Fellow of NAFEMS (National Agency for Finite Element Methods and Standards).

David Fullwood
Professor, Mechanical Engineering
Brigham Young University

With a Ph.D. on "the topology of space-time," David Fullwood is generally interested in the structure of materials and applied mathematical techniques to better understand them. After receiving his Ph.D. from The University of London in 1992, Fullwood spent 12 years in the nuclear industry, transferring centrifuge technology to energy storage flywheels. Critical areas of the mechanical design that he oversaw included hydrodynamic bearings, multifunctional composites, and high-speed rotational behavior. Following the hands-on engineering training, Fullwood returned to academia in 2004 by completing a Master of Science in mechanical engineering. His subsequent academic career began at Drexel University, followed by the past 12 years at Brigham Young University. Fullwood's interests have included microstructure-sensitive design, high-resolution EBSD, and nano-composite sensors.

Krishna Garikipati
Professor, Mechanical Engineering and Mathematics
Director, Michigan Institute for Computational Discovery & Engineering
University of Michigan

Krishna Garikipati obtained his bachelor's degree from the Indian Institute of Technology, Bombay, in 1991, and a master's and Ph.D. from Stanford University in 1992 and 1996, respectively. After a few years of post-doctoral work, he joined the faculty at the University of Michigan in 2000, where since 2012, he has been a professor in the Departments of Mechanical Engineering and of Mathematics. Garikipati's research draws on applied mathematics and numerical methods to explain phenomena in biophysics and materials physics. A recent interest is in using data-driven methods to enhance our ability to solve computational physics problems. In 2016, Garikipati was appointed the director of the Michigan Institute for Computational Discovery and Engineering (MICDE), a research institute focused on developing new paradigms of computational science that cut across application areas. He has been awarded the DOE Early Career Award, the Presidential Early Career Award for Scientists and Engineers, and a Humboldt Research Fellowship.

Marisol Koslowski
Professor of Mechanical Engineering
Purdue University

Marisol Koslowski is a professor of mechanical engineering at Purdue University. She received her Bachelor of Science in physics in 1997 from the University of Buenos Aires, Argentina, and her Master of Science and Ph.D. in aeronautics in 1999 and 2003 from the California Institute of Technology. She was a technical staff member in the Theoretical Division at Los Alamos National Laboratory before joining Purdue. Koslowski has made novel contributions in the development of theoretical models and computational tools to study the mechanical, thermal, and chemical response of materials. The research areas in her group cover a large array of interests, including crystalline plasticity, solid-state transformations, chemical reactions, and fracture dynamics.

Koslowski's group research has been supported by various federal agencies, including grants from the Department of Energy (DOE), National Nuclear Safety Administration (NNSA), National Science Foundation (NSF), Air Force Office of Scientific Research (AFOSR), Air Force Research Laboratory (AFRL), and Office of Naval Research (ONR), as well as by funds from aeronautical (Boeing Company), electronic (Intel Corporation, SRC), and pharmaceutical industries (GlaxoSmithKline and Hoffman-La Roche). These efforts support an interdisciplinary research program with impact in nano and microelectronics, drug and food processing, fiber reinforced polymer composites, nanostructured materials, and energetic materials.

Sankaran Mahadevan
John R. Murray Sr. Professor of Engineering, Department of Civil and Environmental Engineering
Vanderbilt University

Sankaran Mahadevan is John R. Murray Sr. Professor, Department of Civil and Environmental Engineering at Vanderbilt University, Nashville, Tennessee, where he has served since 1988. He has a secondary appointment as professor of mechanical engineering, and directs the Risk, Reliability and Resilience Engineering graduate studies. Mahadevan's research interests are in reliability and uncertainty analysis, machine learning, model validation, material degradation, structural health monitoring, design and manufacturing optimization, and system resilience, with applications to civil, mechanical, and aerospace systems. His research has been funded by NSF, NASA, FAA, DOE, DOT, DOD, NIST, NRC, Airbus, Northrop Grumman, General Motors, Chrysler, Mitsubishi, Union Pacific, and the Sandia, Los Alamos, Idaho, and Oak Ridge National Laboratories.

Mahadevan has co-authored over 600 technical publications, including two books and 300 peer-reviewed journal papers. He is currently the managing editor of ASCE-ASME *Journal of Risk and Uncertainty in Engineering Systems* (*Part B: Mechanical Engineering*), associate editor for two other journals (*ASCE Engineering Mechanics*, *ASTM Smart and Sustainable Manufacturing*), and co-chair of the Prognostics and Health Management Conference (Nashville, 2020). Mahadevan is a Fellow of the American Institute of Aeronautics & Astronautics, Engineering Mechanics Institute (ASCE), and Prognostics & Health Management Society.

Arif Masud
Professor of Engineering Mechanics, Department of Civil and Environmental Engineering and Department of Aerospace Engineering
University of Illinois at Urbana-Champaign

Arif Masud is Professor of Engineering Mechanics in the Department of Civil and Environmental Engineering and the Department of Aerospace Engineering at the University of Illinois at Urbana-Champaign. He received his PhD in Computational Mechanics from Stanford University in 1993. Masud has made fundamental and pioneering contributions to the development of Stabilized and Variational Multiscale Methods for fluid and solid mechanics, residual-based turbulence models, non-Newtonian fluids, and mixture theory based thermo-chemo-mechanical modeling of multiphase materials. He is past chair of the Fluid Mechanics Committee of ASME and past chair of the Computational Mechanics Committee of ASCE. Masud has served as an associate editor (AE) of the ASME *Journal of Applied Mechanics*, AE of the ASCE *Journal of Engineering Mechanics*, and currently serves as AE of the *International Journal of Multiscale Computational Engineering*. He was co-chair of the Finite Elements in Flow Problems Conference (FEF 2019) and is serving as co-chair for the U.S. National Congress on Computational Mechanics (USNCCM 2021). Masud is an Associate Fellow of AIAA and Fellow of USACM, IACM, AAM, ASME, EMI (ASCE), and SES. He has been awarded the 2019 G.I. Taylor Medal by the Society of Engineering Science.

David L. McDowell
Regents' Professor and Carter N. Paden, Jr. Distinguished Chair in Metals Processing
Georgia Institute of Technology

Regents' Professor and Carter N. Paden, Jr. Distinguished Chair in Metals Processing, Dave McDowell joined Georgia Institute of Technology (Georgia Tech) in 1983 and holds appointments in both the GWW School of Mechanical Engineering and the School of Materials Science and Engineering. McDowell served as Director of the Mechanical Properties Research Laboratory from 1992–2012. Since 2012, he has directed the Institute for Materials (http://www.materials.gatech.edu), a Georgia Tech Interdisciplinary Research Institute charged with cultivating cross-cutting collaborations (user facilities, research proposals, workshops) and approaches to accelerate materials discovery and development. McDowell's research focuses on the development of physically based, microstructure-sensitive constitutive models for nonlinear and time-dependent behavior of materials, with emphasis on wrought and cast metals. Topics of interest include microstructure-sensitive computational approaches to fatigue of alloys, atomistic and coarse-grained atomistic simulations of dislocations, multiscale modeling, and systems-based materials design under uncertainty. He has participated in prior TMS study groups and workshops and is a member of the TMS Materials Innovation Committee. McDowell currently serves on the editorial boards of *Nature Partner Journals (npj): Computational Materials* and the *Journal of Multiscale Modeling*, and as co-Editor of the *International Journal of Fatigue*.

Paris Perdikaris
Assistant Professor, Department of Mechanical Engineering and Applied Mechanics
University of Pennsylvania

Paris Perdikaris is an Assistant Professor in the Department of Mechanical Engineering and Applied Mechanics at the University of Pennsylvania. He received his PhD in Applied Mathematics at Brown University in 2015. Prior to joining Penn in 2018, Perdikaris was a postdoctoral researcher at the department of Mechanical Engineering at the Massachusetts Institute of Technology, working on physics-informed machine learning and design optimization under uncertainty. His work spans a wide range of areas in computational science and engineering, with a particular focus on the analysis and design of complex physical and biological systems using machine learning, stochastic modeling, computational mechanics, and high-performance computing. Current research thrusts include physics-informed machine learning, uncertainty quantification in deep learning, engineering design optimization, and data-driven non-invasive medical diagnostics. Perdikaris' work and service have received several distinctions, including the DOE Early Career Award (2018), the AFOSR Young Investigator Award (2019), and the Ford Motor Company Award for Faculty Advising (2020).

Celia Reina
William K. Gemmill Term Assistant Professor, Mechanical Engineering and Applied Mechanics
University of Pennsylvania

Celia Reina is the William K. Gemmill Term Assistant Professor in Mechanical Engineering and Applied Mechanics at the University of Pennsylvania. She joined in 2014 after holding the Lawrence Postdoctoral Fellowship at Lawrence Livermore National Laboratory and the HCM postdoctoral Fellowship at the Hausdorff Center of Mathematics in Bonn, Germany. Reina received her Ph.D. in Aerospace Engineering in 2011 from the California Institute of Technology, following a Bachelor of Science in mechanical engineering from the University of Seville in Spain, and a Master in Structural Dynamics from Ecole Centrale in Paris, France. Among many other awards, Reina is the 2017 recipient of the Eshelby Mechanics Award for Young Faculty, and she currently serves as the recording secretary of the Applied Mechanics Division of the ASME.

William Rider
Distinguished Member of Technical Staff
Sandia National Laboratories

William Rider joined Sandia National Laboratories in 2007 after working at Los Alamos National Laboratory for nearly 18 years. During that time, he has authored well over 100 research papers and two books. Rider's first book focused on high-resolution numerical methods for low-speed and incompressible flows. The second book was a contributed volume on the topic of large eddy simulation. Currently, he works on method and code development for shock physics simulation as well as verification and validation of large-scale simulations.

Rider began his career at Los Alamos National Laboratory, starting in the Nuclear Reactor Analysis Group. There, he developed computer codes to model the thermal-hydraulic behavior of nuclear reactors, including space reactors, and conducted the safety analysis of nuclear reactors. Later, Rider worked on a DOE "Grand Challenge" research project, simulating combustion. While working on that project, his research began focusing on numerical methods for simulating interfaces and interface physics. Later, Rider moved to the Applied Theoretical Physics division, expanding his research interests to include radiation transport, and turbulence computation and modeling. Other research thrusts have included nonlinearly converged methods for simulating radiation diffusion using multigrid Newton-Krylov methods. More recently, Rider's efforts have focused on computational methods for the motion of material interfaces, shock physics, and turbulent mixing. His principal interests are computational physics with an emphasis on fluid dynamics, radiation transport, verification and validation of computational models, turbulent mixing, and models for turbulence.

Rider received his Ph.D. in nuclear engineering from the University of New Mexico in 1992, while working as a staff member at Los Alamos. His Ph.D. thesis was on the design of high-resolution shock capturing methods.

Kiran N. Solanki
Associate Professor of Mechanical Engineering
Arizona State University

Kiran N. Solanki is an associate professor in the School of Engineering of Matter, Transport and Energy at Arizona State University (ASU). Prior to coming to ASU, he was an associate director for the Center for Advanced Vehicular Systems at Mississippi State University. Solanki received his doctorate from Mississippi State University in December 2008. His research interest is focused on characterizing and quantifying process-structure property-relationships across multiple length scales, with emphases that span from fundamental physics to manufacturing of advanced multifunctional materials for extreme applications, including radiation, high rate, fatigue, and creep.

Solanki has co-authored more than 80 journal articles, four book chapters, and 35 conference proceedings. His awards include the SAE Henry O. Fuch Award by the SAE Fatigue Design and Evaluation Committee, TMS Light Metals Magnesium Best Fundamental Research Paper Award, the 2013 TMS Light Metals Division Young Leader Professional Development Award, the 2013 Air Force Office of Scientific Research Young Investigator Research Award, and the 2013 ASME "Orr Award" for Early Career Excellence in Fatigue, Fracture, and Creep.

Michael Tonks
Associate Professor, Materials Science & Engineering / Nuclear Engineering
University of Florida

Michael Tonks is an associate professor of materials science and engineering at the University of Florida (UF). Prior to joining UF in the fall of 2017, he was an assistant professor of nuclear engineering at the Pennsylvania State University for two years and a staff scientist in the Fuels Modeling and Simulation Department at Idaho National Laboratory for six years. Tonks was the original creator of the mesoscale fuel performance tool MARMOT and led its development for five years. He helped to pioneer the approach taken in the Nuclear Energy Advanced Modeling and Simulation (NEAMS) program to use multiscale modeling and simulation (both atomistic and mesoscale) to inform the development of materials models for the BISON tool that are based on microstructure rather than burn-up, and he won the NEAMS Excellence Award for that work in 2014 and the Presidential Early Career Award for Scientists and Engineers in 2017. His research uses mesoscale modeling and simulation coupled with experimental data to investigate the impact of irradiation induced microstructure evolution on material performance. Tonks is also investigating and applying advanced methods for verification and validation with statistical uncertainty quantification.

Recommended Practices Team

The recommended practices team was responsible for development of much of the content of section III of this study report. They participated in several virtual, facilitated workshops, and helped to write, edit, and review section III of this report. Their efforts are gratefully acknowledged, and this team included:

- **William Rider** (Recommended Practices Team Chair), Sandia National Laboratories
- **Mark Carroll,** Federal-Mogul Powertrain (a division of Tenneco)
- **Joshua Kaizer,** U.S. Nuclear Regulatory Commission
- **Jonathan Madison,** Sandia National Laboratories
- **Michele Manuel,** University of Florida
- **David McDowell,** Georgia Institute of Technology
- **William Oberkampf,** Virginia Polytechnic Institute and State University
- **Ravi Shankar,** Scientific Forming Technologies Corp
- **Kiran Solanki,** Arizona State University
- **Barna Szabo,** Engineering Software Research and Development, Inc.
- **Michael Tonks,** University of Florida
- **Brandon Wilson,** Los Alamos National Laboratory

Expert Contributor Satellite Meeting Team

This team met for an in-person facilitated meeting in San Diego, CA, on February 25, 2020, at the TMS Annual Meeting and Exhibition, to provide inputs particularly in areas of this report concerning the verification and validation value proposition, challenge problems, recommended practices, and connecting of experimentalists with modelers. Their efforts are gratefully acknowledged, and this team included:

- **Michael Groeber,** The Ohio State University
- **Tom Kozmel,** QuesTek Innovations, LLC
- **Jonathan Madison,** Sandia National Laboratories
- **Matthew Miller,** Cornell University
- **Amit Pandey,** MicroTesting Solutions
- **David Rowenhorst,** U.S. Naval Research Laboratory

Final Report Review Team

Following the writing and editing of multiple drafts of this report by the TMS science and engineering staff and the lead and recommended practices study teams, an independent review team provided valuable, detailed comments and suggestions for incorporation into the final report. Their efforts are gratefully acknowledged, and this team included:

- **Brad Boyce**, Sandia National Laboratories
- **Wei Chen,** Northwestern University
- **Charles Fisher,** Naval Surface Warfare Center
- **Jacob Hochhalter,** University of Utah
- **Paul Mason,** Thermo-Calc Software Inc
- **Daniel Nicolella,** Southwest Research Institute
- **Zhigang Suo,** Harvard University
- **Xuanhe Zhao,** Massachusetts Institute of Technology

Other Key Contributors

This report was sponsored by the National Science Foundation (NSF), Division of Civil, Mechanical, and Manufacturing Innovation (CMMI), Mechanics of Materials and Structures (MOMS) program, under the direction of NSF program director Siddiq Qidwai, and executed by The Minerals, Metals & Materials Society (TMS). The principal investigator of this activity was George Spanos, Director of New Initiatives, Science, and Engineering at TMS. The technical experts were convened, and their input was compiled into the final report through the efforts of TMS and Nexight Group, LLC. Other important TMS staff contributions to this report include those of Michael Rawlings (Science and Engineering Lead), Amy DeFilippo (Technical Project Administrator), Hui An "Annie" Ooi (Science and Engineering Intern), Bob Demmler (Graphic Designer), David Rasel (Media Manager), Marleen Schrader (Accounting and Human Resources Specialist), Ann Ritchie (Technical Communications Specialist), Lynne Robinson (Department Head, Strategic Communications and Outreach), and Matt Baker (Department Head, Content). Nexight Group, LLC staff members who were heavily involved in this effort include Ross Brindle (Chief Executive Officer) and Jared Kosters (Senior Technical Consultant).

Preface

Who should read this report?

This report is intended to provide value to scientists, engineers, software developers, designers, analysts, regulators, students, and other stakeholders associated with (or intending to work with) computational models related to the mechanics of materials and structures (MOMS). This includes both modelers and experimentalists within the materials science and engineering, mechanical engineering, solid mechanics, structural dynamics, and related communities, spanning academic, industrial, and government affiliation sectors. This report was written with two types of people in mind: novices who have little or no prior experience in robust verification and validation (V&V) and associated/inseparable uncertainty quantification (UQ) practices, and those who have some V&V/UQ experience, but want to establish more rigorous practices. More specifically, researchers, developers, and students associated with materials (both structural and soft materials) and solid mechanics modeling, who utilize advanced computation, materials data, and/or experimental validation tools, should find the information in this report especially useful. It is critical that the community widely adopts robust V&V/UQ practices in order to improve trust, reduce risk, and improve the reliability of MOMS computational models. Beyond practitioners in this field, other stakeholders who can influence the future of advanced computational modeling associated with MOMS should find this report useful, as well. This includes individuals who support financial and/or time investments in science and technologies surrounding computational modeling, such as funding officers and other decision-makers at federal agencies, and leaders/managers in industry. Educators teaching undergraduate and graduate courses related to MOMS, as well as department heads and/or deans within the relevant disciplines, also could use the information in this report to advance associated curricula and enhance research products.

Some guides to navigating this report

Readers are encouraged to navigate this science and technology accelerator study report by first examining the Executive Summary to get an overview of the overall structure and highlights of this document. It is our hope that this report will inspire you to take specific actions consistent with your skills and interests, in order to support development and widespread adoption of robust V&V/UQ practices in models involving the mechanics of materials and structures. The Background, Motivation, and Study Process section (section I) sets the context of this report and summarizes the study goals and process. The Value Proposition (section II) articulates the benefits of the implementation of such V&V and UQ practices. Section III, discussing Recommended Practices for V&V, is arguably the most important section of this report for practitioners working with, or planning to work with, computational models associated with MOMS. It is, in effect, a "field manual" that provides a V&V framework and step-by-step activities within the framework. For planners, decision makers, and policy makers reading this report, section IV provides recommendations on specific benchmark challenge workshops and future funding programs that will help accelerate more robust and widespread V&V/UQ practices in MOMS-related communities. Section V recommends future studies, workshops, symposia, and/or conferences, and section VI further elucidates how modelers and experimentalists might more effectively work together in developing and implementing V&V/UQ methodologies. Finally, some detailed recommendations for improving multidisciplinary training and curricula development in this area are provided in section VII.

As you explore these sections, we aim to draw your focus toward both the tactical details and strategic planning that resonate most with your interests and expertise, and encourage you to prioritize some specific actions that you and your colleagues might undertake immediately or within the next year. Our desire is to provide a significant positive impact to your work, career, organization, and broader technical/professional community.

Call to Action: What actions should be taken after reading this report?

A primary goal of this science and technology accelerator study report is to stimulate direct action by a wide variety of stakeholders who read it, in support of the development and widespread adoption of robust V&V/UQ practices in computational models of the mechanics of materials and structures. Such activity is sorely needed in order to improve confidence and achieve greater efficacy in computational models and to take advantage of their potential value for accelerating innovations in the development of new materials, components, and structures. After reading this report, some general next steps could include: (1) identifying specific V&V and UQ activities that would be most relevant to your personal and/or organizational goals and activities, (2) sketching out a detailed action plan and timeline for you and your colleagues to address the activities discussed herein, and (3) taking concrete steps to initiate these activities. These steps would differ depending on your role and area(s) of interest.

Our desire is that readers will act both promptly and in a sustained, long-term fashion on the recommendations in this report. The specific recommendations and activities identified herein should not be viewed as all-inclusive but used to initiate conversations that determine what would be appropriate for you and your organization. The potential is great for a wide variety of stakeholders who read this report to make rapid progress, as well as foundational, longer-term contributions, toward implementing more robust verification, validation, and uncertainty quantification practices in computational models of the mechanics of materials and structures. Such activity is vitally needed in order to bring to fruition the great potential of these predictive models in supporting the development of advanced new materials, components, and structures.


Months
Months

Executive Summary

Background, Motivation, and Study Process

Computational models are increasingly utilized to guide engineering decision-making, affecting the performance, safety, and longevity of our technology infrastructure. However, despite dramatic increases in the functionality and sophistication of these models over recent decades, such models are rarely sufficiently tested to yield suitably accurate, quantitative results. Particularly for models associated with the mechanics of materials and structures (MOMS), it is imperative that robust verification and validation (V&V) and uncertainty quantification (UQ) practices be widely adopted, since inadequate models and/or misinterpretation of model limits can lead to production delays, unexpected catastrophic failure, costly redesign, and even loss of life.

Some very succinct definitions of verification, validation, and uncertainty quantification are provided at the beginning of this report (section I), as follows:[1,2]

- **Verification** is the process of determining that a computational model accurately represents the underlying mathematical model and its solution.
- **Validation** is the process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model.
- **Uncertainty Quantification** is the process of quantifying uncertainties associated with model calculations of physical Quantities of Interests (QOIs), with the goals of accounting for all sources of uncertainty and quantifying the contributions of the specific sources to the overall uncertainty.

With this background in mind, the overarching goal of this study and report is to provide a framework, detailed recommendations, and resources/references to help accelerate the widespread implementation of robust V&V/UQ techniques and practices within the MOMS and related communities. This study report builds upon the foundation of some key, high-level recommendations provided in a previous TMS workshop and report.[3]

To execute this overarching goal, an internationally recognized lead team of twelve experts, drawing on multiple technical backgrounds and professional sectors, met in multiple, professionally facilitated, in-person and virtual meetings and worked remotely on the content development, writing, and editing of the final report. An equally expert team was convened to focus strictly on the robust recommended practices section of this report (section III), and a satellite meeting with other experts also was held to further explore key concepts initially discussed by the lead team. The content of this report was generated from the efforts described above, then integrated and distilled into an initial draft, which was iteratively edited by the lead study team, the recommended practices team, and an independent final review team of additional experts. Then, all copy editing, design, graphics, and production of the report were completed. (See the Acknowledgements section for specific names and affiliations of the individuals who participated in this study.)

Value Proposition of V&V

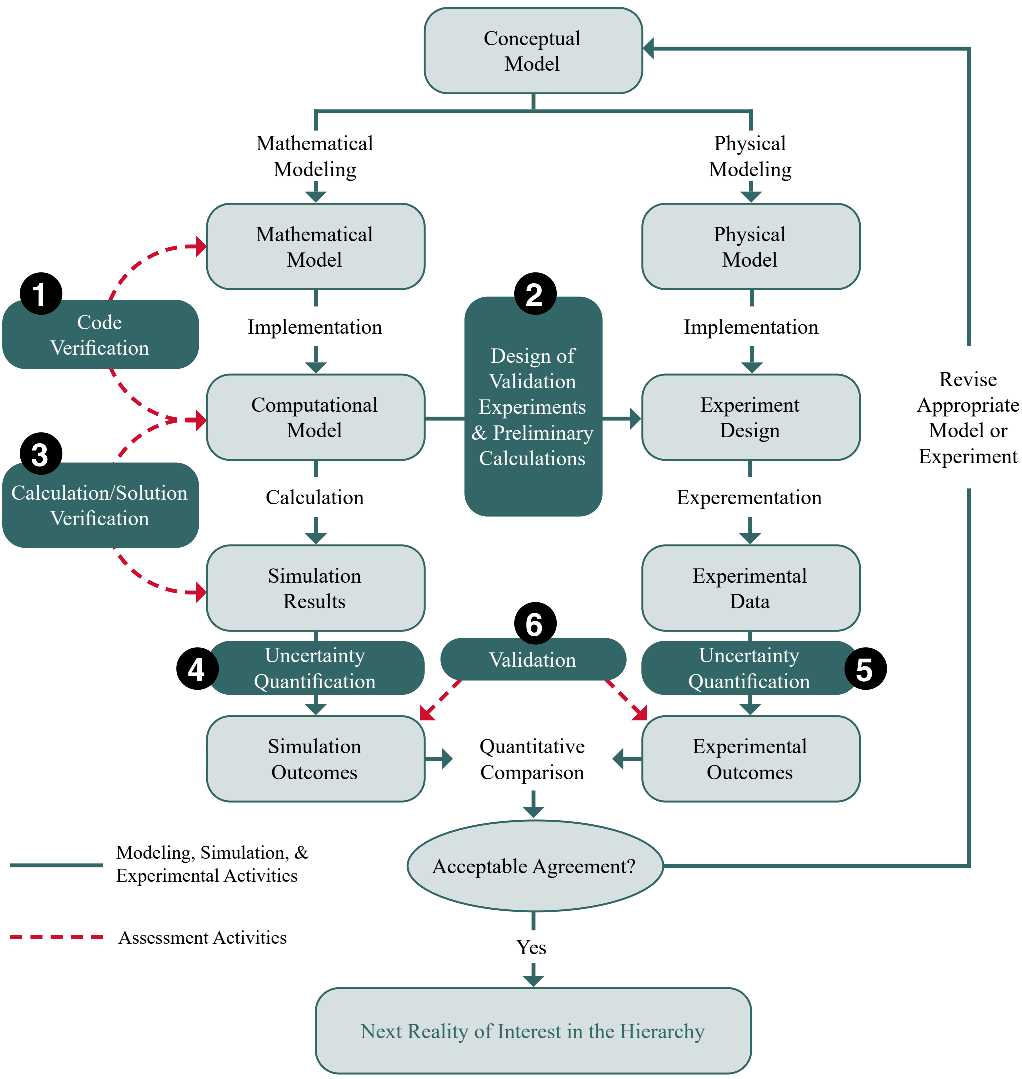
A list of ways that implementing V&V practices can bring value to key stakeholder groups is provided in Table 1 of the Value Proposition section (section II) of this report and is reproduced below. Each of these value elements, including specific justifications and compelling examples, is discussed in detail in section II.

Reproduction of Table 1: Summary table of the value proposition elements for V&V (and associated stakeholder groups).

VALUE ELEMENT	Industry	Government Labs	Academia	Regulators
Faster time to market and increased profitability	X			
Supports risk/failure mitigation and understanding	X	X		X
Better prediction of product quality, costs, and life	X			
Reduces physical testing	X	X	X	X
Improves governmental decision making		X		X
Needed when full-scale tests are not feasible	X	X		X
Assesses and elevates the value of artificial intelligence techniques	X	X	X	
Improves the value and versatility of predictive models	X	X	X	X
Supports assessment of safety and reliability for complex systems	X	X		X
Accelerates regulatory process approval	X	X		X
Enhances the utility and reproducibility of shared data	X	X	X	X

Recommended Practices

Section III on Recommended Practices for V&V and UQ of Computational Models Associated with the Mechanics of Materials and Structures is by far the most robust section of this report and is intended to stimulate implementation of robust V&V/UQ practices within MOMS and related communities. An overarching framework is provided in figure 1 of section III and is reproduced below.



Reproduction of *Figure 1: Simplified V&V flow chart. Numbered boxes represent key V&V process steps. Modified from ASME V&V 10-2006 (R2016).*

In section III, approximately five to ten recommended tasks are listed for each of the six major process steps shown in figure 1. An example of one of those key task charts (for process step # 3 in Figure 1) is reproduced below.

Example of Key Tasks - Calculation/Solution Verification	Contributors
1. Conduct a convergence study.	M, U/CA
2. Calculate a converging sequence of solutions	CD, M, U/CA
3. Subject usable solutions and QoI to a post-calculation analysis	U/CA
4. Subject post-calculation analysis to quality check and peer review	CD, M, U/CA
5. Examine error magnitude for QoI on the baseline mesh	C, U/CA

The suggested community stakeholders/contributors for the recommended tasks described in section III is reproduced below.

Reproduction of Table 2: Community Stakeholders/Contributors involved in the key tasks associated with the different V&V process steps.	
CD	Code/Software Developer
E	Experimentalist
FO	Funding Organization/Resource Gatekeeper
M	Modeler
NA	Numerical Analyst
R/C	Regulator and/or Customer
U/CA	User of the model and/or computational/simulation analyst

For each of the recommended tasks (within each of the six process steps in Fig. 1), methodologies and expected roles of various stakeholders in accomplishing such tasks are discussed in detail, and for readers interested in going deeper, more in-depth references also are provided throughout section III.

Predictive capability of a computational model - in terms of the credibility of how much the model is capable of predicting - is also thoroughly discussed in section III, including the fundamental distinction between the concepts of "validation" and "predictive capability."

Challenge Problems and Sustainable Funding Programs

Domain-specific challenge problem events that promote community building and knowledge transfer across interdisciplinary teams and incorporate significant V&V/UQ efforts will contribute to helping benchmark the efficacy of computational models and developing more robust V&V/UQ practices.

Such challenge problems specifically related to MOMS modeling were thus identified and elaborated upon (section IV) in the following technical domains:

1. Additive Manufacturing - Variability in Mechanical Properties
2. Hardness and Residual Stress in Joined Parts
3. Crystal Plasticity Modeling
4. Interfacial Friction
5. Precipitation Strengthened Alloys
6. Multi-Step Material Modeling

The readers of this report are invited to think of other such challenge problems across a broad spectrum of technical domains and materials (e.g., including both structural and soft materials), related to MOMS modeling.

To incorporate V&V/UQ practices into established scientific workflows (as well as to sponsor challenge problems), it is necessary to obtain sustainable funding. Some past and/or current funding programs are reviewed in this regard, and recommendations are made for specific future program areas to support independent V&V/UQ research programs. Example areas are:

- National Defense-Related Challenge Problems
- Materials Genome Initiative
- Technology Transfer
- Artificial Intelligence

These areas are by no means all-inclusive, as many other (currently funded) arenas could also benefit from such V&V/UQ activities.

Symposia, Conferences, Workshops, and Studies

Suggestions for possible symposia, conferences, workshops, and studies that could contribute to accelerating the adoption of more robust V&V/UQ practices in the MOMS community are discussed in section V of this report. Ideas are considered for symposia/conferences involving a variety of professional societies that could bring value to this topic. A number of other organizations, including some regulatory bodies, government agencies, and relevant conference organizing bodies, also are recommended for involvement in such events. Further recommendations are then provided for developing other, domain-specific, science and technology accelerator studies and workshops that would further advance acceleration of robust V&V adoption within the MOMS community.

Strategies to Connect Modelers and Experimentalists

In order to execute V&V/UQ practices, strong collaboration between modelers and experimentalists is essential. After considering some of the challenges or barriers associated with bringing modelers and experimentalists together, detailed strategies, opportunities, and tactics are discussed in section VI, and summarized in Table 3 (reproduced below).

Reproduction of Table 3: Strategies and opportunities to bring together multidisciplinary teams of modelers and experimentalists.

Strategies	Opportunities/Tactics
Provide Incentive Mechanisms that Motivate V&V Contributions	• Multidisciplinary V&V funding streams and/or requirements • Support for publication venues that reinforce joint experimental-computational V&V efforts • Buy-in of collaborative V&V among industry managers • Robust experimentation requirements for challenge problems
Build Awareness of the Value of V&V/UQ	• Value-driven collaborative V&V/UQ partnerships • Professional society V&V committees/working groups • Recognition and achievement awards for V&V/UQ efforts • Advertisements for shared V&V/UQ resources
Reduce the Entry Barrier for New V&V/UQ Users	• Networking events for V&V/UQ team matchmaking • Joint educational opportunities for experimentalist-modeler teams

Strategies for Improving Multidisciplinary Training and Curricula Development

In section VII, seven possible mechanisms, or strategies, are recommended for building the skill sets needed to conduct proper V&V in computational models associated with MOMS, including:

- Integration of V&V/UQ modules into existing core university courses
- Creation of new core and/or technical elective courses
- Stand-alone tutorials, short courses, and/or workshops
- Open source tools development and dissemination
- Challenge problems
- Tutorials and texts
- Mentoring

Detailed recommendations and tactics are provided for how to address each of these strategies.

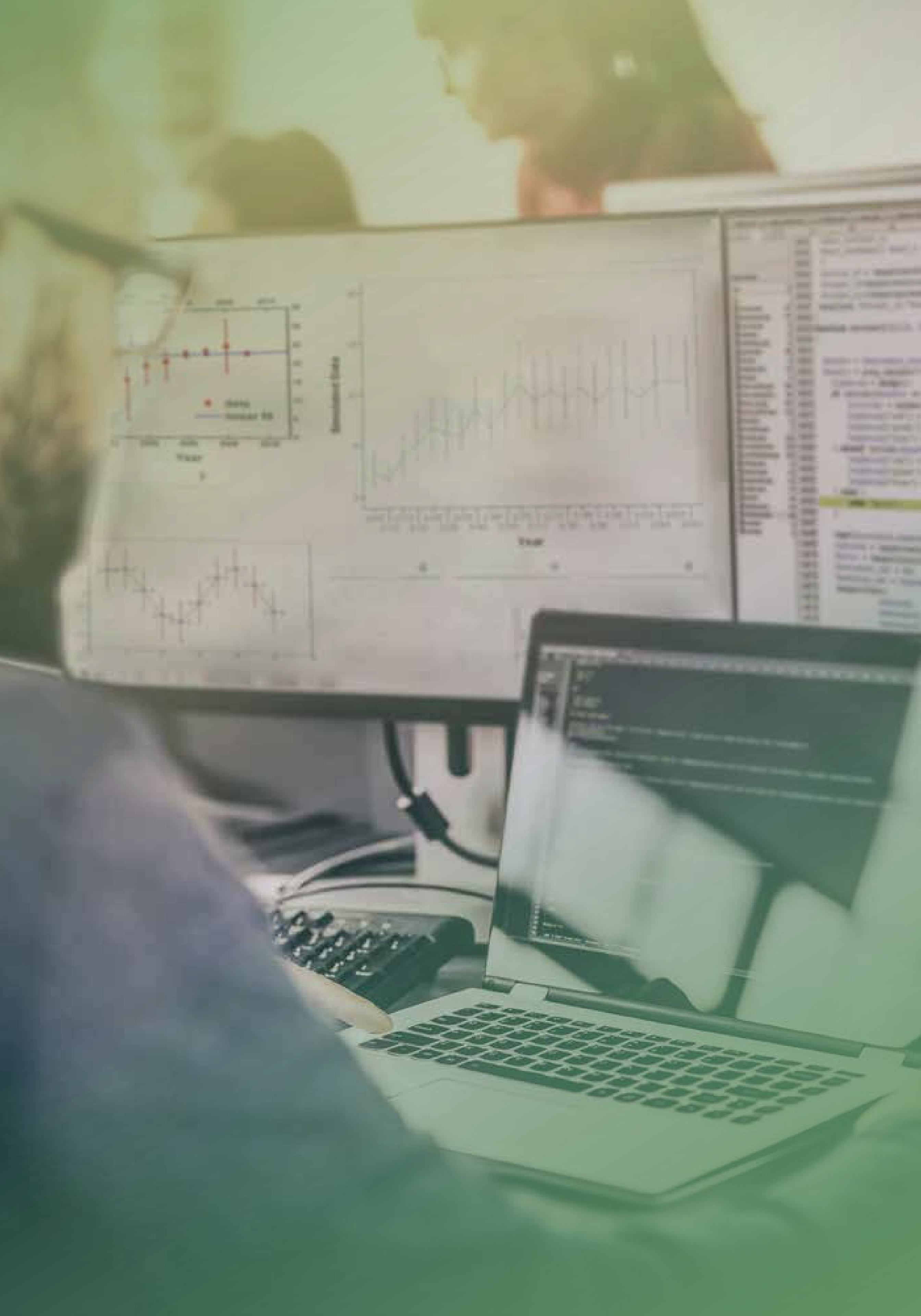
Final Comments and Call to Action

More than just educating the community, a major intent of this report is to stimulate readers to initiate activities and/or practices that encourage implementation of more robust V&V/UQ practices in computational models of MOMS. Another major outcome of this report is clarification of the value added by V&V/UQ activities when conducting all forms of computational simulation. The specific recommendations and activities should not, however, be viewed as all-inclusive, and in that regard, also may be used to initiate conversations to determine other appropriate activities for you and your specific organization. In this vein, section VIII provides further discussion on how to use this report to strongly impact the community going forward.

The potential is great for a wide variety of stakeholders who read this report to make rapid progress (as well as foundational, longer-term contributions) toward implementing robust V&V/UQ practices in MOMS-related computational models. Such activity is vital in order to bring to fruition the great potential of these predictive models in supporting the development of advanced new materials, components, and structures. This progress can draw upon the extensive V&V/UQ experiences and expertise gained and proven in other, related technical arenas.

Glossary

The glossary in the Appendix of this report is more than the typical collection of definitions. It includes detailed descriptions and practical examples of key terms, phrases, and practices discussed throughout this report. Readers at all levels of experience should find this section useful.



I. Background, Motivation, and Study Process

Over the last seven decades, the computer has experienced exponential, worldwide growth, evolving from a few thousand niche devices to a near-ubiquitous industry, which is now intimately intertwined within modern society. More recently, the emergence of computer modeling and prediction of physical behavior and engineering systems has been heralded in the scientific community as one of the most important developments in recorded history;[4] this technology can guide critical decisions that affect every facet of human existence. With the meteoric rise of computing power and capacity, the expectations regarding the use of predictive computational simulations in decision-making have grown proportionately. Today, as we stand at the dawn of a revolution in artificial intelligence, there is immense potential to utilize computer-based predictions to generate critical quantitative information on innovations, new products, activities, and events which could greatly influence the economy, security, and health of people, companies, and nations worldwide. With these wide-reaching, potential benefits comes the immense responsibility to assure the credibility of predictions used for important societal applications.

In accordance with this larger trend, predictive computational models associated with the mechanics of materials and structures (MOMS) offer great potential to significantly reduce the cost, time, and risk to develop new structures, materials systems, and manufacturing technologies.[5,6] While there has been a considerable increase in model sophistication over the last several decades, it is unfortunately quite rare that these models are proven to yield sufficiently accurate, quantitative results for which the level of uncertainty has been adequately quantified. This has led a large segment of the community, in particular industry and regulated sectors, to be reluctant to fully implement and leverage said computational tools due to legitimate concerns over whether the resulting model predictions can be adequately trusted in practical, real-world applications.

This concern is best addressed through the widespread implementation of robust model verification and validation (V&V) and uncertainty quantification (UQ) practices. V&V/UQ is thus an essential part of the model development process, and its importance cannot be overestimated.[7]

To offer useful context, succinct definitions of verification, validation, and uncertainty quantification, summarized from the American Society of Mechanical Engineers (ASME)[1,8,9] and the National Research Council report on *Assessing the Reliability of Complex Models: Mathematical and Statistical Foundations of Verification, Validation, and Uncertainty Quantification*[2] are restated here:

- **Verification** is the process of determining that a computational model accurately represents the underlying mathematical model and its solution.
- **Validation** is the process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model.
- **Uncertainty Quantification** is the process of quantifying uncertainties associated with model calculations of physical Quantities of Interests (QOIs), with the goals of accounting for all sources of uncertainty and quantifying the contributions of the specific sources to the overall uncertainty.

Models can be thought of as tools that allow the complexity of a given system to be reduced to a subset of its essential elements, for instance, in the form of a system of nonlinear, partial differential equations based on mathematical methods relevant to a given field. Simulations then refer to the solution of these equations computationally for conditions that elude analytical solutions. Models and simulations are inherently imprecise because of the use of necessary simplifications and assumptions in models and numerical approximations in simulations. These simplifications and approximations introduce uncertainty into the predictions of the models and simulations. Additional uncertainty is introduced due to the manufacturing process, natural variability in the material, initial processing conditions, wear or damage in the system, and other general system circumstances.[10] As stated by Roy, et al., "each of these different sources of uncertainty must be estimated and included in order to estimate the total uncertainty in a simulation"[10]; this UQ is a vital component of any rigorous V&V activity. Moreover, deep understanding of the magnitude and composition of the various sources of uncertainty allows practitioners and decision-makers to best manage and reduce the risks associated with simulation results in the most efficient and cost-effective way. Without the clear enunciation of the total estimated uncertainties associated with a set of simulated predictions, a decision-maker can easily misinterpret the modeled capabilities of the system in question. As a result, decision-makers could unwittingly waste vast amounts of time and money, as well as potentially put the health and safety of their employees, customers, and companies at risk.

The use of rigorous V&V practices has been shown to expose sources of error and provide for more trustworthy utilization of predictive modeling results when compared to simulations that lacked V&V.[11–14] Moreover, several scientific communities already have a relatively long history of leveraging V&V/UQ techniques; this provides a foundation on which the MOMS community can build. Examples of such related disciplines can be found in the detailed V&V investigation and implementation work of researchers in electrical engineering and computer science;[7] fluid mechanics/dynamics;[15,16] the validation, uncertainty estimation, and/or optimization community;[15,16] and other technical domains, including computational engineering, physics, and/or soft materials within the medical field.[17] Researchers in these areas have covered topics, such as paradigms that relate V&V to the model development process and recommended procedures for model validation;[7] practical frameworks, research needs, and management issues in V&V;[15] a review of V&V in computational fluid dynamics (CFD);[16] and some fundamental issues in V&V such as code verification versus solution verification, model validation versus solution validation, and the distinction between error and uncertainty.[16] In many of these fields, the established V&V framework is applied near-ubiquitously to modeling efforts across both spatial and temporal length scales; committing to a similar practice will be critical to successful adoption of a V&V framework within the materials community, as it utilizes models at a wide variety of length scales.[18]

The computational models and V&V frameworks in some of these domains are typically of greater maturity compared to those within the contemporary material science and solid mechanics communities. This is, in large part, because these disciplines deemed a "systematic, rigorous, and disciplined approach to verify and validate computational models and methods...essential to confirming the accuracy of model predictions."[19] As a result, these fields have repeatedly engaged in community-wide initiatives that promote the identification and implementation of model/simulations verification and validation guidelines.[1,8,9,19] Thus, it is critical for the MOMS community to more broadly accept that computational model V&V/UQ practices and techniques are equally essential to their work and adopt similar community-wide initiatives to promote their implementation in a robust way.

Study Goals and Process

The major purpose of this study and report is to provide the motivation, framework, recommendations, and resources to accelerate the widespread implementation of robust V&V/UQ techniques and practices within MOMS and related communities. This study report builds upon several key recommendations and opportunities within the MOMS community for stimulating the adoption of rigorous V&V approaches, which were first identified within the *Verification & Validation of Computational Models Associated with the Mechanics of Materials* workshop and report.[3] Six of the recommendations, which were deemed of the highest priority and/or most promising from that report, comprise the central thesis of this study and are examined in depth in the subsequent sections (see section III - section X).

These six high-priority recommendations[3] are:

1. Enunciate a strong value proposition for widespread V&V in solid mechanics, materials science and engineering (MS&E), and related communities
2. Develop a state-of-the-art V&V recommended practices document to guide practitioners
3. Define new benchmark/challenge problems and sustainable funding programs for collaborative V&V projects
4. Develop action plans for specific technical society sessions, conferences, and/or workshops addressing V&V strategies in focused subdomains
5. Outline strategies for connecting modelers and experimentalists
6. Suggest strategies for incorporating V&V concepts into multidisciplinary training and curricula.

During the development of the present study, a core team of twelve internationally recognized experts was convened by The Minerals, Metals & Materials Society (TMS) via professionally facilitated in-person workshops, online meetings, homework assignments, and a series of virtual/remote workshops, all of which generated much of the content of this report. A separate team also convened to specifically develop the recommended practices section of this report. Additionally, a smaller, satellite team met briefly to provide further, specific contributions. All of these outputs were integrated and distilled into a draft of this final report, which was iteratively edited by the lead study team, the recommended practices team, and an independent final review team of experts (all listed in the Acknowledgements section). All copy editing, design, graphics, and production of the final report were then completed.

II. Value Proposition of V&V

A strong V&V/UQ value proposition is essential, as in many cases, significant resources and cultural "buy-in" will be required to stimulate meaningful, widespread, and robust V&V/UQ. Due to the extensive utilization of computational methods across various sectors throughout the solid mechanics and materials communities, independent value propositions may need to be developed for each of the various stakeholder groups, since their motivations can vary widely. A list of some overarching ways in which implementing V&V practices can bring value to key stakeholder groups is shown in the Value Element column of Table 1, with the subsequent columns indicating the stakeholder group(s) that should significantly benefit from each value element.

Because they are so intricately intertwined, it is often difficult to distinguish between the value proposition of general computational simulations and the value of applying V&V to simulations. In this vein, models and simulations are tools that approximate a given system or process, while the application of robust V&V practices is an organized accumulation of evidence to quantify the accuracy of such models and simulation results, thereby allowing decision-makers to more confidently rely on simulation results.

For the purposes of this discussion, the value proposition elements reflect the potential impact of the rigorous application of V&V/UQ practices in various employment sectors across the mechanics and materials communities. The general premise is that the value of the simulation decreases with diminishing V&V practices due to a lack of demonstrable model confidence and trustworthiness.

Table 1: Summary table of the value proposition elements for V&V (and associated stakeholder groups).				
VALUE ELEMENT	Industry	Government Labs	Academia	Regulators
Faster time to market and increased profitability	X			
Supports risk/failure mitigation and understanding	X	X		X
Better prediction of product quality, costs, and life	X			
Reduces physical testing	X	X	X	X
Improves governmental decision making		X		X
Needed when full-scale tests are not feasible	X	X		X
Assesses and elevates the value of artificial intelligence techniques	X	X	X	
Improves the value and versatility of predictive models	X	X	X	X
Supports assessment of safety and reliability for complex systems	X	X		X
Accelerates regulatory process approval	X	X		X
Enhances the utility and reproducibility of shared data	X	X	X	X

Faster time to market and increased profitability

Although developing the proper tools and infrastructure for verification and validation can involve a significant investment in time and resources, when done correctly, V&V offers the potential to accelerate the rate of introducing innovative products to market in a more cost-effective and reliable manner. For example, in the practice known as Integrated Computational Materials Engineering (ICME),[20,21] case studies such as Ford Motor Company's virtual aluminum castings and QuesTek's development of aircraft landing gear have shown that ICME has resulted in acceleration by as much as five years to get products qualified and/or to market, saving development costs in excess of tens of millions of dollars.[21]

More specifically, such computational modeling - along with proper V&V/UQ practices to ensure sufficiently reliable models - can greatly reduce costs and development time by significantly decreasing the number of characterization experiments and property/performance tests. Increased confidence in simulation results is needed for decision making in areas including, but not limited to, alloy chemistry selection, thermomechanical processing, and part manufacturing processes.[21]

Supports risk/failure mitigation and understanding

Validation and UQ activities can be considered an accumulation of evidence which quantifies the accuracy of and confidence in an underlying simulation. These resulting "evidence packages" are vital for decision makers to have confidence in simulation outputs and are used in several sectors, such as aircraft and nuclear industries, to mitigate design failures. In the field of gas turbine engines for commercial aircraft, computational simulation, combined with extensive operating data on fan, compressor, and turbine blades, has helped to nearly eliminate the occurrence of blade failures during flight.

In nuclear power plant operations, simulation of diverse failure scenarios, along with regular inspection and maintenance, has reduced the likelihood of serious plant failures in normal operations to near zero. Additionally, simulations provide a very high degree of confidence that even if those unlikely scenarios occurred, the plant would remain in a safe state.

Better prediction of product quality, costs, and life

Broadly, companies that employ rigorous V&V/UQ practices on models used to predict downstream costs, new opportunity areas, new products, and/or product life have a competitive advantage over those who do not adequately verify and validate their models. The process of verification and validation invariably inspires model improvements as various shortcomings are exposed. Companies that are early V&V/UQ adopters are more likely to produce timely and accurate model results, leading to superior products and/or systems. In this fast-paced industrial landscape with worldwide competition, better trust in the implementation and decision support gained from these models is critical for both practitioners and managers to make confident decisions in order to maintain cutting-edge technology and competitive advantages.

An example of such a decision encompasses the terms and duration of a product warranty. Often when a physical product is developed and sold, a warranty is provided for a specified timeframe within which the company is liable for repairs and replacements. Companies expend large amounts of resources to calculate appropriate durations that best limit their risk. If this is done through untrustworthy or inaccurate methods, the results can be costly in terms of money, reputation, and potentially dire legal consequences. In such cases, the development of V&V/UQ evidence packages is an effective way to simultaneously garner confidence in a final warranty decision and mitigate substantial financial risk.

Reduces physical testing

Especially for complex systems, the cost to perform the volume of iteratively designed experiments and tests necessary to converge on an optimized design and performance for the individual components, as well as for the overall system, is prohibitive. The use of computational models and simulations can vastly reduce the amount of physical testing required, but their use represents a significant risk if the implementation is not verified properly or if models/simulations do not accurately reflect real-world phenomena. Robust V&V/UQ defines the degree to which the computational results are dependable enough to sufficiently reduce the amount of physical testing required. It also shifts the focus of the physical testing away from defining the overall physical behavior and toward validating the computational models and simulations defining the physical behavior in small regions of high uncertainty.

Improves governmental decision making

V&V/UQ produces critical evidence that inspires confidence for faster, more informed decision-making pertinent to the missions of national laboratories, government agencies, and regulatory authorities. As almost every U.S. government laboratory is funded, at least in part, by an entity within the executive branch, all have a specific mission that informs the scope of its research and development activities. These entities represent a huge portion of scientific research in the United States and are constantly at the forefront of new technological innovations, designed for rapid, practical implementation.

As the utilization of computational tools within the critical areas of national defense and energy increases, the use of V&V/UQ practices should increase proportionally in these research and development efforts to allow for both faster execution of the mission and an attendant increased confidence in the results. V&V/UQ also provides evidence to show where capabilities are lacking or are primed to inform investment portfolios and research priorities. In addition to giving decision-makers assurance in the simulation results, V&V/UQ evidence packages provide a structure to easily and reliably convey that confidence to colleagues and superiors up the chain of command when the work is mission critical. Moreover, in highly regulated fields, such as nuclear power and medical devices, the regulatory authority appropriately requires considerable V&V/UQ evidence before trusting simulation results. The added confidence, reliability, and accessibility provided by V&V/UQ practices to the execution of such missions is an absolute necessity.

Needed when full-scale tests are not feasible

Even in large-scale research and development facilities, such as national laboratories, resources to replicate full-scale, real-life conditions are not available or affordable, especially when the required testing systems or equipment are quite large and complex. In other cases, full-scale testing is legally or ethically prohibited or is functionally impossible. Moreover, the equipment required for a legacy, full-scale, physical test may be antiquated or unavailable under current policy and/or environmental requirements. Additionally, time and resource constraints can preclude such full-scale, real-life test capacity. More generally, real-life testing can be simulated, saving untold amounts of money and time. This is useful as long as the simulated results can be trusted to be an appropriate proxy for experimental testing. In each of these cases, though, simulations help decisions and allow testing to go forward, only so long as rigorous V&V/UQ programs are in place to assure the quality and credibility of the simulations.

Assesses and elevates the value of artificial intelligence techniques

Artificial intelligence (AI) can be defined, broadly, as the field of computer science that studies how non-sentient agents learn from their environments. Its subdiscipline, machine learning, provides the opportunity for computer-assisted learning from data. AI and, more specifically, machine learning have enjoyed an explosion of interest in engineering, reflecting similar trends in science, as well as in business, entertainment, retail, sports, politics, and virtually every field of human endeavor. In what follows, the term AI is used, but in most cases it is machine learning methods that are actually in play.

AI opens doors to expanding human intellectual endeavors and deepening the understanding gained in scientific work. It allows such work to be filtered through a lens that is separate from and additive to human abilities. However, as the rate of discovery has accelerated and more of the underlying understanding of AI "black box" predictions is outsourced, there is a danger of becoming blind to the need for more consistent, independent review of these processes. With the emerging ubiquity of AI, it is vital to have a rigorous way of evaluating the validity of these outcomes, and in order to improve confidence and trust in AI-derived outcomes, practitioners must develop and standardize the rigorous requirements of the V&V/UQ process within their scientific methodology. The exponential growth that AI is experiencing across nearly all sectors, disciplines, and industries makes it difficult to overstate the need for a systematic V&V/UQ process to assess the reliability of AI results.

Improves the value and versatility of predictive models

Similar to the use of AI, the use of predictive models is quite appealing for advancing scientific and engineering frontiers. One difference is that predictive models are still based on mathematical methods inputted by a practitioner (rather than on data input into the computer itself) and are verified and validated using existing V&V/UQ techniques. If said V&V/UQ techniques are performed on model results for a strategic sampling of boundary conditions within a larger area of interest, a quantitative metric of the model's reliability from the resulting V&V evidence packages would provide high confidence in the model when it is utilized to make predictions at other, non-V&Ved boundary conditions within the area of interest. Having quantitative metrics for assessing prediction results at previously unexplored conditions would not only allow for rapid exploration of vast amounts of unknown landscape in a variety of arenas, but it also could create standards by which models can be assessed by funders and the greater scientific community. In this vein, in terms of scientific publications and patents, the US is among the world leaders in scientific innovation,[22] but beyond these coarse metrics, it is often difficult to measure the impact of our scientific investments.

V&V/UQ is seen by many as an extension of the scientific method to be applied in the computer age, allowing predictive computational modeling to be trusted to deliver the expected performance benefits or advantages of a new application or product of interest. As scientific and economic enterprises use more computational modeling tools and techniques, V&V/UQ will provide increased intellectual rigor as well as an alternative metric to quantitatively evaluate the economic and intellectual value and utilization of such models. In other words, the intellectual rigor provided by V&V/UQ is important to assure the quality and impact of such computational modeling results. Additionally, V&V/UQ defines the limits and bounds of applicability, providing evidence of where research investments can most profitably yield expanded predictive capabilities.

Supports assessment of safety and reliability for complex systems

For complex systems, V&V/UQ provides increased confidence in the simulation of a multiplicity of scenarios. Misleading or inaccurate simulations of system failures, if they occur, can be devastating to companies, society, or the environment. Ancillary decisions about health and safety spending for long-term employee, environmental, and business health, also are made with greater confidence for these complex systems when based on rigorous V&V/UQ.

Accelerates regulatory process approval

In heavily regulated industries, it is very time consuming and arduous to get any new material or system design approved for use. For example, although a 10 percent failure rate may be acceptable in an academic research endeavor, for the Federal Aviation Administration (FAA) considering airplane engines, one in ten engine failures would be unacceptable. Thus, regulatory agencies are extremely cautious when approving innovative technologies and, therefore, rely almost exclusively on physical test results in lieu of simulated outputs. However, such approval historically requires many years of iterative experimentation and millions of dollars in the process. The national stockpile stewardship program, on the other hand, has effectively replaced full-scale testing of nuclear weapons with a number of tools where the crucial contributor is computational simulation with V&V/UQ.[23] The utilization of V&V/UQ is a key element to providing the necessary confidence in the computational results. The possibility of including simulation outputs and V&V/UQ evidence packages as viable components of a regulatory approval package can vastly speed up time to market of any given material, while still upholding an agency's standards of reliability. As previously indicated, the danger of excluding V&V/UQ on computational models or using it in a perfunctory way is that decision makers, in this case regulators, will not have confidence in the ability of simulations to provide adequately and appropriately trustworthy results.

Enhances the utility and reproducibility of shared data

In today's interconnected economy, companies, universities, and government laboratories must communicate data internally across divisions and externally across the globe amongst customers, colleagues, and practitioners. In all cases, reliable data should be findable, accessible, interoperable, reusable (FAIR), and trustworthy by recipients and/or users; however, at this point in time, shared research results are lacking in this regard. Project and personnel turnover are contributors to these problems, but robust application of V&V/UQ during the innovation process could minimize the impact of such turnover. V&V/UQ must be a critical part of the technical information being communicated, so that one can better reproduce and rely on what was done by others. More importantly, V&V/UQ will enable data to be reinterpreted and built upon quickly, saving substantial costs in terms of both time and financial resources.

III. Recommended Practices for V&V and UQ of Computational Models Associated with the Mechanics of Materials and Structures

Building off recommendations in the previous TMS workshop study report on V&V,[3] it was recognized that an important step in stimulating widespread implementation of V&V and UQ practices in MOMS computational models - within the materials science and engineering (MS&E), solid mechanics, and related communities - was to develop a recommended practices document. The intent of this section is to serve as that document, which includes recommended practices to address basic elements of V&V/UQ within the MOMS domain and could equally be applied to other technical domains, as well. These basic elements are built on successful approaches from other technical sub-disciplines. This section can also serve as a starting point for those interested in going into more depth when applying V&V/UQ to their problems. Note that there are many publications that provide much greater detail and robust guidance on this topic, and the reader is herewith referred to some of these relevant publications.[1,2,7,8,15,16,24–28]

As alluded to in the earlier report[3], this section can also be thought of as "...a 'field manual' of sorts for modelers, experimentalists, and other domain specialists working in this area, including both those that have some experience with V&V and those that have little or no experience," and "at the very least, it should help raise awareness of robust V&V practices in the solid mechanics, and materials science and engineering communities."[3] The recommendations in this section should in no way be thought of as all-inclusive, since every computational model will have its own set of nuances, restrictions, and complexities, i.e., there is no "one-size-fits-all" V&V/UQ solution for all models and/or application domains. Instead, this section should be thought of as a first step or example template with some guiding principles that can serve as a foundation for assessing a specific problem of interest.

In a related vein, it is not expected that all detailed tasks recommended throughout this section will be undertaken by all practitioners implementing elements of this report. Instead, each individual will need to take away as much as they are able to use, based on their specific computational modeling-related activities, as it is recognized that readers will be at various levels of experience, expertise, and engagement in V&V/UQ practices.

The recommendations in this section should in no way be thought of as all-inclusive, since every computational model will have its own set of nuances, restrictions, and complexities, i.e., there is no "one-size-fits-all" V&V/UQ solution for all models and/or application domains. Instead, this section should be thought of as a first step or example template with some guiding principles that can serve as a foundation for assessing a specific problem of interest.

In a related vein, it is not expected that all detailed tasks recommended throughout this section will be undertaken by all practitioners implementing elements of this report. Instead, each individual will need to take away as much as they are able to use, based on their specific computational modeling-related activities, as it is recognized that readers will be at various levels of experience, expertise, and engagement in V&V/UQ practices.

III.1. Framework, Definitions, and Overview

Figure 1 (below) provides a graphical representation or flowchart that serves as a foundation from which to conceptualize V&V steps and to develop some of the recommended practices. It was modified from a V&V guidance document published by the American Society of Mechanical Engineers (ASME) - ASME V&V 10-2006 (R2016).[24]

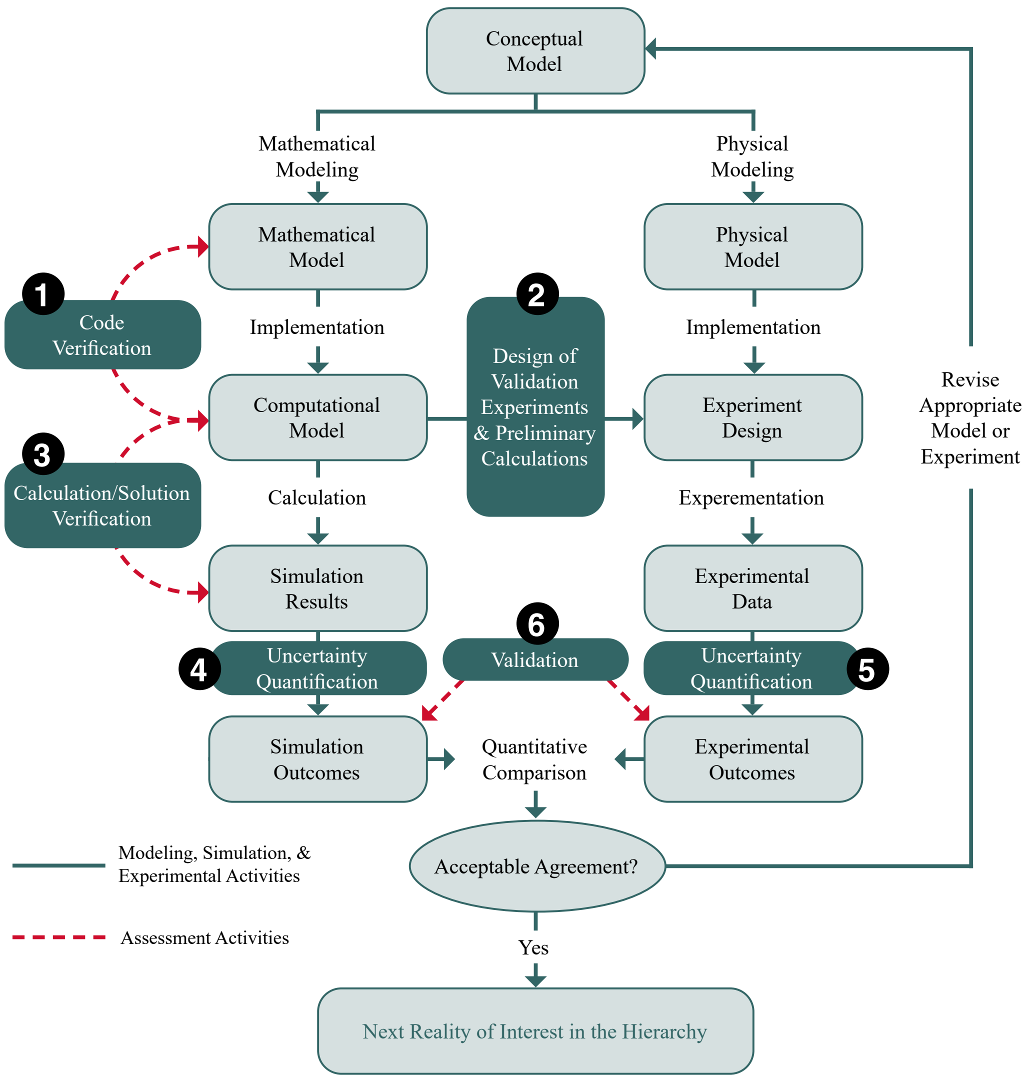


Figure 1: Simplified V&V flow chart. Numbered boxes represent key V&V process steps. Modified and adapted from ASME V&V 10-2006 (R2016).[24]

Similar to the introductory discussion at the beginning of section III, this flowchart (Fig. 1) is intended to serve only as a guideline or general framework from which to build V&V and UQ practices for a given complex problem, recognizing that there is not a one-size-fits-all solution to V&V/UQ for every application of each computational model. V&V is, instead, an iterative process, repeated multiple times as the model is developed and implemented for specific Quantities of Interest (QoI).

QoI can be described as physical entities or features related to the target outputs and prediction goals of the mathematical model, the values of which are of interest to stakeholders, typically because they inform decisions. QoI can be experimentally measured quantities, model input quantities, or system response quantities of interest. Examples of QoI from previous modeling efforts include, but are not limited to: (1) thermal conductivity of a microstructure; (2) maximum Von Mises stress experienced in a part; (3) the ratio between the maximum temperature in a part and the materials melting temperature; (4) the difference between the maximum tensile stress in a part and the maximum allowable stress.

The V&V process typically begins with individual subsystems and connections across subsystems, i.e., how they interact, and often links one scale (spatial or temporal) to inform the next hierarchical scale. There is significant and inherent uncertainty associated with such linking across subsystems, as well as from one scale to the next. Moreover, as new technology is developed, corresponding models are validated in different ways, depending on the "technology readiness level" (TRL), which is a common system to vet the maturity of the technology on a scale of 1 to 9. In addition to lab-scale validation of the models, which is typical at TRL 4 or 5, and is represented by box #6 in Figure 1, there are additional steps to show that the model still works adequately for real-world events (TRL 6 and 7), that the model is deployed for field use (TRL 8), and that the model works in the field (TRL 9).The flowchart in Figure 1 does not explore the validation steps for technologies at TRL 6 or higher. Neither does it express the concept of the model application domain (see section III.3). Despite these words of caution and prudence, the subject matter experts involved in this study believe Figure 1 is a very useful, generalized foundation/framework from which practitioners and other stakeholders can build strong V&V/UQ practices for computational modeling, related to the MOMS community, as well as to other technical domains.

In section III.2 below, detailed recommended practices are provided, and a set of specific tasks are outlined and discussed for each of the process steps (numbered boxes) in Figure 1. However, before considering those detailed recommended practices tasks, it is important to consider each of the process steps at a higher/synopsized level. For that reason, provided immediately below for each process step is: (1) an overarching description, including some key elements; (2) currently "acceptable" V&V/UQ practices for MOMS-related computational modeling; and (3) a brief overview of the more robust recommended V&V practices that should be adopted (which are considered in much more detail in section III.2).

1. Code Verification

Description
Code verification refers to the process of determining whether computer code accurately represents an underlying discretized, computational model, i.e., whether the code solves the modeling equations correctly. A method of doing this compares the software output of the computational model to exact/known solutions. In other words, code verification addresses the solution accuracy produced by the mathematical calculations performed on a computer, as compared to the exact result. It should also be used to ensure that the discrete convergence of all of the algorithms and the software steps used to produce the computed solution align with the theoretical values for the specific spatial and temporal discretization schemes.

Current Practices
Some current practices for code verification that are generally acceptable in the MOMS community include occasionally comparing software results to a known analytic (or trusted) solution and qualitative determination that the software delivers some expected behavior. There also are a wide variety of software quality control practices related to testing, bug fixing, model verification, and version control.

Recommended Practices
More robust practices are needed and should include: (1) software quality assurance (SQA) approaches using automated testing against known solutions on a regular basis, e.g., regression testing, which is carried out whenever a change is made to the code base; (2) employing known solutions based on established analytical or manufactured solutions;[29–31] (3) regular but less frequent, e.g., annual, comparison of spatial and temporal convergence rates with theoretical values; (4) quantitative error estimates, e.g., using the L2 norm for FEM errors [32]; (5) consistent and established protocols by software developers for code implementation, testing and verification; (6) end users verifying the results for select problems representing their specific application. See the *Code Verificatio*n subsection of section III.2 for a robust discussion of specific tasks and details of these practices.

2. Design of Validation Experiments and Preliminary Calculations

Description
An important goal of this step is to promote, at an early stage of the V&V process, improved communications between the modelers, computational analysts, and the experimentalists designing and conducting the validation experiments. This will result in experiments that are more effective in assessing the accuracy of computational models.

Current Practices
In current MOMS-related communities, there is typically little, if any, communication between computational analysts and experimentalists conducting validation experiments. Moreover, experimental measurement uncertainties and/or input data uncertainties are rarely rigorously investigated or reported. Consequently, the detailed experimental information needed for the mathematical model is commonly not reported in the validation experiment documentation.

This generally leads to preliminary calculation steps that lack proper quality control and UQ, in which case the validity of the mathematical model is insufficient with regard to the most important physical phenomena of interest.

Recommended Practices

As described in detail in the recommendations discussed in section III.2, more robust practices should include: (1) computational practitioners informing experimentalists in advance of the type of mathematical model being used and the computational model assumptions, approximations, and input data required; (2) examination of specific QoI within the computational model for expected numerical behavior and rejection of QoI that are singular or ill-behaved by their nature; (3) experimentalists informing the modelers/computational analysts of experimental procedures and diagnostics that are planned, the range of conditions that are possible, and the uncertainties in the experimental measurements; (4) the computation experts and experimentalists reaching agreement on what input data can be directly measured (as well as that which should be calibrated using inverse solutions of the mathematical model), accounting for both experimental and modeling uncertainties during preliminary calculations; (5) initial and boundary conditions in the experimental configuration should be measured, understood, and addressed in the computational model. These practices and collaborations also can unveil details and assumptions that would otherwise be accepted without as much attention and might confound UQ.

3. Calculation/Solution Verification

Description

Solution verification generally refers to the process of determining how well the discretized model represents the mathematical model, based on QoI. Unlike code verification, an exact solution is not known in solution verification, and thus cannot be applied in this process. Perhaps a better way to conceptualize solution verification is that it involves estimation of the errors of approximation in terms of the computed quantities of interest. This approximation (or calculation) error is defined as the difference between the estimated fully convergent solution and the numerical solution for the discretized computational model. Solution verification can, therefore, also be considered to be an evaluation of the reliability of the numerical solution with error estimation, via numerical estimation of calculation error, as a function of discretization features in the computer software.

These concepts will be more fully elucidated when considering the detailed tasks and recommendations for calculation/solution verification described in Section III.2 and are considered in detail in other references, as well.[27,33–36] As a general rule, though, the goal of solution verification is to verify that errors of approximation are negligibly small in comparison with model form errors and experimental uncertainty.

Current Practices

In many cases, current solution verification practices are encompassed by one or more of the following: limited or no error estimation or different mesh discretizations are performed; assertion of mesh independence/convergence of solutions is made without appropriate quantitative evidence to support this assertion, e.g., often just an appeal to visual inspection of the solution is made, and/or global quantities, such as global error norms, are substituted for more rigorous analyses.

Recommended Practices

More robust practices that should become more broadly accepted and applied by the MOMS community include: (1) performance of numerical error estimations on all system response QoI; (2) use of mesh refinement to estimate discretization error in the model solution and the estimation of the mesh converged solution (the meshes should ideally be self-similar); (3) use of fitting techniques to estimate the mesh convergence rate for the extrapolated solution with negligible discretization error; (4) examination of the dependence of the solution on other numerical choices, such as the iterative solver tolerance, the finite element choice, or the order of approximation.

4. Uncertainty Quantification (UQ) - Simulation

Description

UQ is a critical component of computational modeling; its importance cannot be overstated.[26] UQ generally refers to quantification of the uncertainty, e.g., changes to the solution, due to various factors influencing the simulation outputs (with the possible exception of quantification of numerical uncertainty, which is evaluated by solution verification). UQ associated with the simulation side of the flowchart in Figure 1 encompasses the act of generating a mathematical description of the uncertainty in specific models and inputs, as well as the act of using those mathematical descriptions within a computer simulation to determine the resulting uncertainty in the QoI. In other words, simulation UQ can be described as the act of quantitatively characterizing the uncertainty for simulated QoI from all contributing uncertainty sources, including simulation input uncertainties, constitutive model form uncertainty, and uncertainty due to QoI post-processing. Simulation UQ typically involves applying mathematically rigorous probabilistic methods to estimate the confidence of a prediction and then quantifying the model accuracy and/or reliability.

Generating a mathematical description of the uncertainty of inputs or models requires intimate knowledge of those inputs and models; the greater the knowledge, the better the mathematical description. In general, three methods are used to mathematically describe the uncertainty of an input or model, from least to greatest information: (1) the bounds approach; (2) mean and variance methods; (3) distribution methods. In the bounds or interval approach, only the upper and lower bounds of the variable are determined and used in the subsequent analysis. In the mean and variance approach, only the mean and variance of the variable are determined and used in the subsequent analysis, e.g., central differencing methods. In the distribution approach, a probability distribution is determined or assigned and used in the subsequent analysis, e.g., Monte Carlo sampling methods. These methodologies are discussed in detail by Wang and McDowell.[26]

Current Practices

Current practices often neglect numerical uncertainties and correlations between inputs. There is limited focus on model-form uncertainty, i.e., the uncertainty due to the assumptions and approximations made in the formulation of the mathematical model of the system of interest. Moreover, there are inadequate or imperfect surrogate models, data sampling schema, discretization, reproducibility, replicability, etc. In most practices, there also is very limited effort applied to UQ related to estimating the discrepancy between predicted model data and experimental data.

Recommended Practices
More robust UQ practices on the simulation side of Figure 1 include: (1) determining the most important uncertainties that should be quantified, e.g., which inputs and which models; (2) determining the best method to mathematically describe each of those uncertainties; (3) based on the mathematical description, choosing an appropriate method to combine the uncertainties into an overall uncertainty of a given QoI; (4) including quantifying numerical uncertainties found via solution verification through the formulation of statistical sub-models, which are calibrated against existing data; (5) robust testing of sampling methods and underlying model emulators; (6) application of inverse methods to calibrate the model using experimental data for bias correction and parameter calibration.[26]

5. Uncertainty Quantification (UQ) - Experimentation

Description
Experimental uncertainties can originate from experimental equipment and facilities, material characterization techniques, diagnostics, physical fluctuations, and/or post-processing approaches. UQ of experimentally based uncertainties is the process of quantitatively characterizing the uncertainty (both random and systematic) for a measured QoI, based on all necessary inputs/ parameters involved in executing an experiment that is employed to approximate physical phenomena associated with the predictions of the simulation.

Current practices
Current practices for quantifying experimental uncertainties span a broad range of activities, varying from low to high quality, including: (1) little to no experimental uncertainty is reported; (2) uncertainties are determined as a set of limits, e.g., +/- 5%, (3) uncertainties at specified confidence levels are quantified using forward Taylor-series error propagation methods and statistical methods, such as ISO/GUM (Guides to the expression of Uncertainty in Measurement).[37–39]

Some deficiencies in current experimental UQ practices include: (1) neglecting variability between facilities, including the sequence of test procedures or environmental conditions that may affect the QoI; (2) uncertainties are often determined from small sample sizes or single tests; (3) comparisons between small lot or single test experiments are often made along with large lot tests, with no accounting for statistical variance expected (i.e.,based on varying sample size or sometimes geometry); (4) post-experimental data processing and analysis that affect uncertainty of the QoI are not fully documented; (5) calibration methodologies and the frequency of those calibrations for experimental equipment are assumed but not reported or documented with the results.

Recommended Practices

It is imperative that more robust experimental UQ practices become widely adopted. Some practices that should be implemented to every extent possible include: (1) sample size of the experimental measurements should always be reported; (2) sample size should be maximized to provide accurate estimates of the underlying QoI variation; (3) uncertainties on QoI and inputs should be characterized by confidence limits, at a minimum, e.g., see ISO/GUM guidelines[37,38]; (4) ideally, joint probability distributions and covariances for QoI and inputs can be measured or characterized; (5) uncertainties from and sensitivities to data post-processing can be propagated into measured QoI using Taylor-series error propagation or Monte Carlo methods; (6) QoI can be measured by multiple diagnostics and/or on multiple experimental facilities to determine systematic errors from the diagnostics and facilities (while quantifying variability between facilities); (7) uncertainties, sensitivities, and correlations between QoI and all simulation inputs, such as facility design and material characterization, can be characterized; (8) any systematic cleaning of data, e.g., elimination of measurement outliers or bias correction for the measured QoI, should be reported; (9) statistical design of experimental methods should be employed in designing experimental test matrices that maximize the value of information extracted regarding QoI and provide more effective support for validation of associated predictive computational models.

6. Validation

Description

In the broadest sense, validation of a computational model refers to the process of determining the degree to which the computational model is an accurate representation of physical reality from the perspective of the intended uses of the model.[1,3] Validation involves the quantitative comparison of calculated results to experimentally measured values. In practice, this involves experimental data that may have been obtained under complex conditions and/or environments, e.g., complex loading conditions, that may not have been used before in the context of validation. It is not appropriate to use experimental data in model validation that has already been utilized to calibrate the model. In certain cases like atomistic modeling, experimental results may be limited, owing to experimental resolution in time and/or space; so, ab initio models might serve as a basis for validation of interatomic potentials. It also should be recognized that in validation practices, there is no such thing as a universally "validated" model; in general, when the term "validated" is used, it should be interpreted as the comparison of the model to empirical data that has been judged to be relatively complete for a particular use case. Section III.3 also addresses, in depth, an important distinction that must be made between "validation" and "predictive capability."

Current Practices

Current practices for validation of computational models within the MOMS community vary widely and are often qualitative or insufficient. They might involve some basic comparisons of calculated results, i.e., code output, such as stress/strain curves or microstructural texture, against more complex experimental results. Generally, the current state of practice is to subjectively characterize the quality of simulation as being "close" to experiments without regard to the computational or experimental uncertainties. For example, if a given level of mesh discretization is found to yield good agreement between computation and experiment, the model is considered "validated." This subjective assessment of quality typically is left to determination by expert opinion.

Recommended Practices

More robust validation practices need to be adopted. Examples include: (1) rigorous numerical testing to confirm whether the model represents the intended physics and physical space/domain; (2) comparison, i.e., quantification of the degree of agreement, of the model with the tests/experiments, which include UQ from both (a) the model (including numerical error and model form UQ) and (b) the tests/experiments; (3) use of data in validation as similar as possible to the real-world application of the computational model, ensuring that the data used for validation has not been previously used in the development or execution of the computational model, i.e., for calibration or training of the model; (4) ensuring that the model and experiments should have similar error tolerance, standard deviation, and confidence levels; (5) understanding important concepts concerning the validation domain, as discussed in detail in section III.3.

III.2. Recommended Practices - Process Steps, Tasks, and Contributors

For each of the process steps (numbered boxes) pictorially represented in Figure 1 and described above, a set of key recommended practices/tasks and the expected roles of various V&V-related community members or stakeholders are presented in each of the sub-sections below. For a given process step, the tasks may not all necessarily be in sequential order according to the way they are numbered. Since many different types of practitioners need to be involved in these recommended key tasks, a starter list of the types of community members/stakeholders that could serve as contributors to these tasks is provided in Table 2.

Table 2: Community Stakeholders/Contributors involved in the key tasks associated with the different V&V process steps.	
CD	Code/Software Developer
E	Experimentalist
FO	Funding Organization/Resource Gatekeeper
M	Modeler
NA	Numerical Analyst
R/C	Regulator and/or Customer
U/CA	User of the model and/or computational/simulation analyst

It should be noted that it is common for a single individual to cover multiple roles described in Table 2. For example, a single person can often serve as CD, M, NA, and U/CA. Similarly, a single person can represent R/C and FO. For larger simulation activities, the roles become more distinct. It is critical that these roles be executed by experienced, detail-oriented individuals, as the potential for human error can have a significant impact on the validity of the results. There is often an order of operations, a knowledge of requisite source data, valid and invalid assumptions, and correct interpretation of the outputs – all of which are required in constructing an evidence document – and all depend, to some extent, on the interpretation of the analysts. In other words, a naïve human operator can make even very good codes yield incorrect results if the analyst fails to perform the analysis and/or interpret the results adequately.

The abbreviations provided in Table 2 are used in the next series of charts, which summarize the recommended key tasks and contributors for each of the V&V process steps. Following each of these process step task tables is a discussion of each of the tasks within these tables.

Similar to the earlier discussion of the V&V practices flowchart and process steps, the recommended tasks also should be viewed only as guidelines or frameworks from which to build V&V and UQ practices, as each computational modeling effort will have its own unique set of circumstances. Not only is each problem different, but each person involved in the computational modeling project/effort may have different levels of expertise and experience in V&V/UQ, ranging from minimal to a vast amount of experience, so that the "Contributors" column also should be viewed as a general guide for the types of individuals who would undertake these tasks. The intent here is that each individual reading this section can find, use, and apply the knowledge and references provided herein in accordance with the level of applicability that meets their needs and experience level.

1. Code Verification

Key Tasks - Code Verification	Contributors
1. Provide funding for (and prioritization of) code verification activities	FO
2. Review code documentation	CD, M
3. Provide a comprehensive examination of the governing equations solved by the computational model	CD, M, NA
4. Provide strategy for numerical method implementation	M, NA, U/CA
5. Define test problems with known exact solutions	M, NA
6. Carry out spatial and temporal convergence studies	CD, M, NA
7. Identify and resolve computation errors in the code	CD, NA
8. Examine sufficiency of the code verification efforts	M
9. Establish an automated system for code verification when changes are made to the code base	CD
10. Document the verification process and build the evidence package	CD, M, NA

Task 1. Provide funding for (and prioritization of) code verification activities

This is an important starting point to support the execution of the remaining recommended tasks associated with code verification. That is, without targeted funding support, the amount of effort that practitioners will invest in code verification associated with computational models of the mechanics of materials and structures (MOMS) could be limited, due to the drain on time and resources associated with the many other aspects of such computational modeling efforts. Yet, code verification is a critical step in the V&V process. Funding organizations or management-level decision-makers who control resource allocation must provide the funding required for the many different activities associated with code verification (see the tasks above). Just some examples of support that could be provided include supplemental funding for computational resources associated with high-performance computing (HPC), and support for the development of manufactured solutions (i.e., exact solutions to governing equations) – see Task 5.

Task 2. Review code documentation

This recommended practice task would likely involve the code/software developer, as well as the person doing the actual computational modeling (the "modeler"). Preferably, in addition to computer science and mechanics expertise, one or both should have an extensive applied math background. For instance, this task could include engagement of code developers to retrieve and communicate workflows, input/output data and formats, and any regression testing that has been incorporated, along with verification and validation procedures that were considered at the time of code development. Furthermore, with the increased tendency to couple codes via application programming interfaces (APIs), all such capabilities and expected type definitions should be explicitly documented. As a specific, computational mechanics example, a detailed description of the included finite element formulations should be made. For example, in legacy finite element codes, one can find elements that were developed before the foundations of the modeling methodology were fully established, and in such cases, at least some of those elements may be laden with what are referred to as "variational crimes," i.e., they fail to satisfy the conditions of consistency and stability and/or reduced integration is used. Thus, it is critical to properly document the finite element formulations.

Task 3. Provide a comprehensive examination of the governing equations solved by the computational model

The objective of this recommended practice task is to assess, in extension to Task 2, the relevance of previously completed verification studies along with the validity of the currently incorporated physics. This task could involve the code developer, the person doing the computational modeling, and/or a numerical analyst. These are just examples of the types of people who might contribute, and it is not a hard and fast rule as to who would be involved in this task, within particular organizations or personnel groups. For example, people with strong backgrounds in numerical analysis who might contribute to this task could be found within code development groups as well as in applied math groups, depending on the organization, its structure, and its personnel composition.

Specifically, this task would involve a mechanics-based description of the input and output parameters of the models as related to the solution of the underlying equations, i.e., examining whether the input data is available and accessible or how it can be acquired. It also should include assessing the relevance of benchmark solutions published by the code developer. If the class of applicable problems specific to the code verification project is not adequately covered by those benchmark solutions, then additional problems or solutions to those equations should be formulated and solved, with focus on the requirements specific to the current computational modeling effort.

Task 4. Provide strategy for numerical method implementation

In addition to a numerical analyst, participants in this effort could include the computational modeler, some other users of the model, and/or a computational/simulation analyst. The numerical analyst's primary focus is the detailed technical, numerical specification of the solution of the governing equations, including the definition of the stability and accuracy of the approximations. The computational analyst, by contrast, is focused on the use of the computer code to solve problems of interest, including the validation cases. This strategy would cover the order of accuracy, applicability, and limits of the numerical method implementation. Numerical schemes should be defined to help inform the code structure as to how the equations are solved. Related to the discussion in Task 2, this task also could address ascertaining that the finite elements employed satisfy the conditions of consistency and stability; it must be ensured that the elements used form a converging sequence when refining the mesh for well-posed problems.

Task 5. Define test problems with known exact solutions

This task might best be accomplished by some combination of the computational modeler and/or a numerical analyst. It would include determination of the specific, exact solutions to be employed for code verification; such solutions could be readily available analytical solutions or could be "manufactured solutions." The method of manufactured solutions (MMS) refers to a general procedure for generating an exact solution for code accuracy verification, and the basic idea is to manufacture an exact solution without being concerned about its physical realism. Details of this method have been provided by various authors [29–31], and one description of a manufactured solution is that it is an exact solution to some partial differential equation (PDE) or set of PDEs that has been constructed by solving the problem backwards.[31]

The specific exact solutions (either analytical or manufactured) should come from a well-established, peer-reviewed source, and thus be grounded in solid mathematical techniques. In the case of computational problems built around the fracture of materials, for instance the Sandia fracture challenge[11–14], these exact solutions could be related to fracture criteria informed by peer-reviewed experimental studies. The persons who define these test problems and exact solutions need to understand the assumptions used to derive the exact solution(s), e.g., a small strain assumption in the simulation. The references and implementation method used for the analytical solutions should be provided, and in the case of manufactured solutions, the MMS should be formulated and/or taken from the literature. Previous references/scientific papers can be used to define the correctness of the implementation of the manufactured or analytical solutions.

Then, the exact solutions are used for comparison to find algorithmic and programming errors in the code, i.e., to build an understanding of the level of accuracy and applicability.

Task 6. Carry out spatial and temporal convergence studies

This task involves quantifying convergence rates and the error in the software outputs, as compared to known solutions, across a sufficient parameter space. Established norms should be employed across the applicable parameter space, e.g., one suitable norm for FEM errors is the L^2 norm.[32] A common strategy is to refine the mesh, compute the solution on the finer mesh, and use the solutions on the two meshes for a qualitative comparison. This task might best be accomplished by some combination of the code developer, the computational modeler, and/or a numerical analyst. These convergence studies should be conducted in whatever discrete degrees of freedom the code utilizes. Typically, this includes a discretization in space through a mesh and time with a time step size for time-dependent problems. The numerical error characteristics of the code in each of these variables are essential to quantify.

Task 7. Identify and resolve computation errors in the code

This should be undertaken by the code developer, perhaps working in coordination with the modeler and/or a numerical analyst. In this case, the project leader (most likely the modeler) should iterate with the software/code developer to fix and improve the code, primarily in the case of open source codes. Common practices for code developers, such as unit and regression testing, should be followed to assure the code is fit for use. Unit testing involves writing functions to test fundamental sections ("units") of the code. Unit tests should cover a significant percentage of the overall code base and should be assembled to test more complex combinations.

Regression testing involves rerunning a suite of such tests over time to assure that code updates, bug fixes, etc. do not break previously working capabilities. Finally, coordinated efforts between the modeler, code developer, and numerical analyst should focus on identifying and testing expected use cases of the entire software framework.

Task 8. Examine sufficiency of the code verification efforts

This task is primarily the responsibility of the modeler. The modeler should make an assessment of whether the physics of interest are covered by the code verification efforts, as well as understand the limits to which the code has been verified. This involves strong communication with software developers to review their evidence packages to confirm verification for the application of interest. For example, verification of a simple problem does not guarantee the software is adequately verified for the more complex, real-world scenarios to which the model is being applied. It is important to recognize that analytical solutions of governing equations are relatively limited in many cases. It is possible that sufficiency is not achievable given these limitations. Under these circumstances, the modeler should ensure that the verification is as complete as reasonably possible.

Task 9. Establish an automated system for code verification when changes are made to the code base

This task is primarily the responsibility of the software developer, when making changes/updates to the code. In such cases, there should be procedures in place for code verification of all updated software. This task should be automated in a manner similar to the unit and regression testing practice described in Task 7. However, in contrast to the relatively simple unit tests, more comprehensive testing of the code with complete verification model definitions is done here and outputs are compared to known, mechanics-based solutions. These automated verification tests should be made part of the regression test suite.

Task 10. Document the verification process and build the evidence package

The "evidence package" in general refers to the evidence of verification and validation procedures. For instance, the evidence package for the verification process includes input and output from the code, analysis of the solutions, and subsequent convergence analysis. This task would primarily involve the code/software developer and the computational modeler (and perhaps a numerical analyst). In particular, the evidence package should be provided to the modeler by the code developer to support the claims of code verification. This applies to commercial off-the-shelf (COTS) code verification conducted by software engineers and/or open source code verification that is well-documented and may be independently verified. The convergence analysis should give a comparison between the observed and expected convergence. The expected convergence is provided by the numerical analyst through theoretical means. In addition, the descriptions of the verification problems should be given or documented through the archival literature.

2. Design of Validation Experiments and Preliminary Calculations

Key Tasks - Design of Validation Experiments and Preliminary Calculations	Contributors
1. Address funding needed to facilitate experimentalist-modeler V&V coordination	FO
2. Determine the quantities of interest (QoI)	CD, U/CA
3. Define the purpose and intent of the experiments	E, M, R/C, U/CA
4. Provide preliminary model inputs and simulations	E, M, U/CA
5. Describe methods for measuring required model input data	E, U/CA
6. Develop and prioritize approaches for measuring QoI	E, U/CA
7. Maintain/preserve records of all associated datasets	E, FO, M, U/CA

Task 1. Address funding needed to facilitate experimentalist-modeler V&V coordination

It is important for officers at funding organizations or other resource gatekeepers in leadership positions in industry, academia, and national laboratories to allocate funding to facilitate ongoing experimentalist-modeler collaborations before, during, and after the experiments. It is critical to have this continuous feedback loop of communication and interaction during the planning stages, as well as the execution and the analysis stages of the experimental efforts. This will both incentivize and subsidize the costs of such critical cooperation.

Task 2. Determine the quantities of interest (QoI)

This is a task that must be accomplished early in the V&V process, and directly relates to process step #3, tasks 3 and 5, described further below. This task might typically involve some combination of users, computational/simulation analysts, and code developers. They would identify the QoI to be studied and/or predicted by the computational model and describe the expected behaviors of these QoI.

Task 3. Define the purpose and intent of the experiments

This is a key primary task that must be accomplished early in the V&V process and can involve coordination among a number of different roles, including experimentalists, computational modelers, users of the model, computational/simulation analysts, associated numerical analysts, and relevant customers and/or regulators. In these early discussions, various tradeoffs should be considered, e.g., limitations of the experimental facility and instrumentation, and funding, time, and schedule constraints.

This task involves determining which physical phenomena the experiment is exploring and providing guidance and recommendations for the appropriateness of such experiments for validation purposes specific to the project simulation efforts, i.e., the use case. The physical phenomena of interest and goals of the validation experiment must be defined up front. During this task, a comprehensive list of assumptions and approximations in the formulation of the mathematical model is provided. It is also important to quantify the initial state of the material, i.e., the values of the relevant state variables. To provide experimental data against which the implemented model and numerical methods can be compared, special care should be taken to ensure that the experimental setup does not violate any of the significant assumptions made in the mathematical model formulation.

In case this cannot be strictly followed, the sensitivity of the QoI with respect to differences between model assumptions and experimental conditions must be quantified.

Task 4. Provide preliminary model inputs and simulations

For this effort, experimentalists work with modelers and computational/simulation analysts involved in the project. Preliminary simulations should be performed to aid in the design of validation experiments. Initial model inputs are provided to aid in the design of the experiment, e.g., the boundary conditions of the model, the sensor placement for capturing data, calibrated material parameters, etc. A list of potential experiments and methods for measuring QoI to be predicted should be developed during this effort. This may encompass a menu of experiments that includes trade-offs relative to the application domain of interest and the outputs of the preliminary simulations. Simulation/computational analysts can explain the input data (boundary conditions, material properties, initial conditions, etc.) needed for the mathematical model, in order to design and select the proper validation experiments.

Task 5. Describe methods for measuring required model input data

Experimentalists should work with model users and/or computational/simulation analysts involved in the project to determine the best methods for experimentally measuring model input data, such as material properties, and/or boundary conditions, such as external load or temperature fields. This task describes and prioritizes approaches for measuring (and if necessary, calibrating) the required model input data.

Statistical sampling tools can be employed to estimate experimental uncertainties in the model input data. This practice can include gage repeatability and reproducibility (GR&R).[40–42] A GR&R study is conducted to obtain replicate measurements on units by several different operators. The gage and part variance components are then estimated by conducting analysis of variance (ANOVA), which has been defined as "…the process used to evaluate a 'gauging instrument's' accuracy by ensuring its measurements are repeatable and reproducible. The process includes taking a series of measurements to confirm that the various output measurements are correctly identified, and that the same measurements are obtained under the same operating conditions over a set duration."
(See https://asq.org/quality-resources/gage-repeatability.)

There are situations where the model input data cannot be directly measured experimentally, e.g., stiffness and damping in joints of assembled structures. For these situations, experimentalists and simulation analysts could come to agreement on a method to calibrate input data using inverse solutions of the mathematical model.

Task 6. Develop and prioritize approaches for measuring QoI

In this task, experimentalists should confer with computational/simulation analysts on approaches for measuring the system response QoI. A list should be developed with potential experiments and methods for measuring the QoI to be predicted. In other words, a menu of experiments should be developed to include any needed trade-offs (prioritization) relative to the specific domains of interest. Capabilities and limitations of the experimental facilities should be considered, relative to the stated goals of the experiments, including limitations and viable options of the proposed experiments. If computationally simulated QoI cannot be measured in the experiment, experimentalists and the computational/simulations analysts should discuss other important response quantities that could be measured in order to set the experimental goals for this step.

Task 7. Maintain/preserve records of all associated datasets
It is essential that all datasets associated with validation experiments and preliminary calculations be captured, archived, and maintained with proper version control and metadata. This requires some financial commitment to maintain the proper data registry in a sustainable fashion, and thus would involve funding organizations and/or other resource gatekeepers coordinating with modelers and experimentalists in support of such data infrastructure. In cases where sufficient data platforms were not already in place or available to users, this also could involve the initial development of such a platform.

3. Calculation/Solution Verification

Key Tasks - Calculation/Solution Verification	Contributors
1. Conduct a convergence study	M, U/CA
2. Calculate a converging sequence of solutions	CD, M, U/CA
3. Subject usable solutions and QoI to post-calculation analysis	U/CA
4. Subject post-calculation analysis to quality check and peer review	CD, M, U/CA
5. Examine error magnitude for QoI on baseline mesh	C, U/CA

Task 1. Conduct a convergence study
A critical part of solution verification is performing a convergence study. The objective is to estimate the limit value of each QoI as the number of degrees of freedom in the computational model is systematically increased. For example, as mesh density is refined by a given ratio, e.g. factors of 2, corresponding changes in QoI are compared to previous refinement steps. The difference between an estimated limit value and the value corresponding to the greatest number of degrees of freedom studied is an estimate of the error in the QoI. Analysts must be mindful that the QoI may or may not converge monotonically. Theoretical limits of the calculation/solution verification should be determined. (As an example, this could address the inability to maintain accuracy for finite rotations or geometric nonlinearity on otherwise linear elasticity models.) The convergence study is best undertaken by some combination of computational modelers, users of the model, and/or computational/simulation analysts.

Task 2. Calculate a converging sequence of solutions
This task follows and/or is intimately related to Task 1, and participants could include code developers, computational modelers, other users of the model, and/or computational analysts. It would involve calculation of solutions on the mesh sequence and extraction of QoI information from this set of calculations. A discretization scheme should be determined in which the numerical discretization, e.g., finite element mesh, is defined, and integration and interpolation formulations are selected to satisfy the conditions of stability and consistency. Exploration can be given to multiple types of refinement schema, e.g., uniform, graded in geometric progression, radical (non-quasi-uniform) meshes with various gradient parameters, etc. A triage of solutions and QoI should be conducted to determine viability of the computed solutions, and computational parameters should be examined for idiosyncrasies in solution behavior, such as solver settings, iterative tolerances, etc. Error estimators also could be incorporated into the codes to indicate if codes are being pushed beyond specified accuracy thresholds.

A sequence of numerical solutions can be obtained by refining the discretization, increasing the interpolation order, or both. For the likely context of FEA-based modeling and simulation, the reader is referred to the following to address the types of efforts described in this task, as well as in tasks 4 and 6 below.[27,33–36]

Task 3. Subject usable solutions and QoI to post-calculation analysis

Computational/simulation analysts could then undertake post-calculation analysis that includes the following elements. Extrapolated QoI and rates of convergence should be examined to define error estimates. The QoI can be plotted against the number of degrees of freedom on a semi-logarithmic scale. Limit values of the QoI can be estimated by extrapolation and then used to estimate discretization errors. The limit values and estimated errors should be reported as part of the calculation/solution verification process.

Task 4. Subject post-calculation analysis to quality check and peer review

Computational/simulation analysts can provide the results of their analyses in Task 3 to modelers, code developers, and/or others as a quality check of the calculations and a method of peer review of these findings.

Task 5. Examine error magnitude for QoI on baseline mesh

In this task, computational/simulation analysts and/or code developers would examine the error magnitude for QoI on the baseline mesh. Then, they would confirm the adequacy of the error level for the analytical requirements and convert the error estimates to uncertainties. These uncertainty estimates would be provided to the UQ team for proper quantification of the uncertainty (see below).

4. Uncertainty Quantification (UQ) - Simulation

Key Tasks - UQ - Simulation	Contributors
1. Characterize/define all uncertainties in the simulation	M
2. Conduct sensitivity analysis on various sources of uncertainty	FO, M
3. Run simulations to generate distributions of solutions	E, M, U/CA
4. Quantify and rank inputs that impact uncertainty of QoI	U/CA
5. Propagate uncertainties from post-processing into measured QoI	M, CD, U/CA
6. Assess error bounds of simulation outcomes	M, CD, U/CA

For further background and details on the various UQ efforts described in the next series of recommended practices tasks, the reader is referred to a number of references that describe these UQ methodologies in more depth.[2,8,26,37,38]

Task 1. Characterize/define all uncertainties in the simulation

To set the foundation for this V&V/UQ process step, the modeling limits, parameters, and general uncertainties first must be described/defined. It is important to distinguish between epistemic and aleatory uncertainties. Aleatory uncertainties refer to inherent randomness or variability in any QoI. Epistemic uncertainties, on the other hand, refer to forms of uncertainty arising from the practitioner's lack of knowledge concerning any QoI, system of interest, or application scenario of interest.

Task 2. Conduct sensitivity analysis on various sources of uncertainty

Although computational modelers would undertake this task, the significant effort expended not only here, but in many of the other tasks, as well, would require support from funding organizations or management entities in charge of allocating resources for the project. In this regard, for high-fidelity/complex models, an assessment of cost implications for new hires, workstreams, and/or high-performance computing requirements would be needed.

This task would involve quantifying the effect of uncertainties from various sources (e.g., model-form, parameter, input) on the simulation QoI; i.e., how do uncertainties in model inputs, parameters, etc., affect uncertainty in model outputs? Model-form uncertainties refer to uncertainty due to assumptions and approximations made in the formulation of the mathematical model. For example, assuming a specific random variable behaves according to a specific empirical equation or is best-represented by a specific type of distribution results in some model-form uncertainty. Parameter uncertainty is generally focused on the uncertainty of specific values that appear in the model, such as the value of a coefficient in an empirical model or the value of the mean or variance in a normal distribution. Input uncertainties are focused on the specific input (or boundary) values, which the model uses to generate a result.[26]

While the sources of uncertainty are considered distinct, it may not be possible to distinguish between their contributions in the simulation results. A supporting activity for this task could be to employ optimization tools to enhance the progression of the UQ activities and make the analytics more efficient and/or economical.

Task 3. Run simulations to generate distributions of solutions

This task involves running simulations to assess the UQ and the distributions of solutions arising from the identified uncertainties. Modelers can provide users with descriptions of the model inputs, QoI, and corresponding uncertainties - model-form, parameter, and input uncertainties. Dataset experimental input uncertainties should be provided in a format needed to estimate the total UQ. Finally, a baseline nominal model test (with nominal inputs) could be run to study model behavior, i.e., use the same code, but with a new model or mesh.

Task 4. Quantify and rank inputs that impact uncertainty of QoI

Inputs that impact the uncertainty of the QoI should be quantified and ranked as to their impact on the simulations. This should be done by the primary users of the modeling code.

Task 5. Propagate uncertainties from post-processing into measured QoI

Modelers, code developers, and/or computational analysts should take responsibility for propagating uncertainties from post-processing into the measured QoI and analyze the sensitivity of critical inputs on variations of QoI.

Task 6. Assess error bounds of simulation outcomes

An assessment should be made of error bounds of simulation outcomes with respect to the imposed parameters. If appropriate experimental data is available, this is where the experimental data (see also process step 5 below) is used to determine the simulation UQ results. This task would best be undertaken by some combination of modelers, code developers, and/or computational analysts.

5. Uncertainty Quantification (UQ) - Experimentation

Key Tasks - UQ - Experimentation	Contributors
1. Identify modeling uncertainties to tailor the set of experiments	M
2. Quantify variability of experimental tests	E, M, R/C, U
3. Ensure differentiation between validation and calibration steps	E, U/CA
4. Ensure sufficient funding/budget for validation experiments	FO
5. Prepare/decide on format for publishing information on experimental uncertainty	E, FO

Task 1. Identify modeling uncertainties to tailor the set of experiments

In this step, a model sensitivity study is conducted to support the design of the proper experiments. In particular, model estimates and/or sensitives to experimental data should be assessed.

Task 2. Quantify variability of experimental tests, i.e., repeatability of experimental results

This task would be performed primarily by experimentalists working in coordination with a combination of modelers, relevant regulators or other customers relying on the model, and/or other model users. Experimental uncertainty thresholds/requirements that will help inform the validation assessment can be defined in this task. At a minimum, the experimentalists must report and document the experimental conditions, sample sizes, calibration methodologies, and systematic errors. To whatever extent possible, they should invest effort into quantifying uncertainties, sensitivities, and correlations between QoI and simulation inputs, based on the diagnostics/measurements, facility designs, and/or material characterization. In particular, assessment and quantification can be made of potential uncertainty and error bounds, i.e., distributions of values, of the experiment itself, and local/specific facility error (uncertainty) in such experimental results - both systematic and random. Recommendations/guidance should be made concerning both the repeatability of the experimental results and improvements in the confidence of the experimental results/outputs. Judgments can be made as to whether the model and experiments have appropriately similar error tolerances, standard deviations, and confidence levels. In this task, it will be important to coordinate with any regulatory agencies that might be involved or have a stake in these predictive simulations to ensure the proper standard for acceptanceis maintained, e.g., repeatability, safety margin.

Task 3. Ensure differentiation between validation and calibration steps

Relative to uncertainty quantification of the experimental data, it is imperative to differentiate clearly between validation experiments and any calibration steps in which experimental data was used. That is, the validation experiments should be completely independent from experimental data that was used for model calibration earlier in the modeling efforts.

Task 4. Ensure sufficient funding/budget for validation experiments

In order to justify the needed funding support for experimental UQ and validation efforts, it will be important to educate funding organizations and/or other resource managers concerning the rationale for using/acquiring (new) software and/or experimental methods in order to ensure funding/budgets are adequate for these efforts.

Task 5. Prepare/decide on format for publishing information on experimental uncertainty

This relates to most of the previously listed tasks and includes identifying sources of uncertainty in the experiments, compiling the experimental UQ data, e.g., error bars, and reviewing any relevant standards, e.g., from ASTM, for publication of such experimental UQ data, as well as communication of such relevant standards to the publisher.

6. Validation

Key Tasks - Validation	Contributors
1. Review boundaries of applicability of the model	M
2. Compare computational outputs to experimental results and their error ranges	E, M, U/CA
3. Quantify validation metrics and produce data visualizations	M
4. Define validation limitations and assumptions	E, M, U/CA
5. Designate funding/resources for developing validation metrics	CD, FO, R/C
6. Document and disseminate validation metrics among all stakeholders	M

Task 1. Review boundaries of applicability of the model

Predominantly, it is the responsibility of modelers to review boundaries of applicability of the model and conduct assessments of model inaccuracies as a function of individual experiments. Any given model has physical limits of applicability where the modeling is valid and accurate. An example would be elasticity, which loses applicability when plastic deformation begins. The same is true for modeling assumptions such as boundary conditions. It is noted, though, that individual experiments may or may not clarify the predictive capability of the model – see section III.3 on "Predictive Capability."

Task 2. Compare computational outputs to the experimental results and their error ranges

This task would be performed by a combination of computational modelers, experimentalists, users of the model, and/or computational/simulation analysts. In this effort, computational outputs would be directly compared to experimental results (and their error ranges) in a quantitative fashion. To whatever extent possible, consideration can be given to the distribution metrics, e.g., mean, standard deviation, kurtosis, skewness, and other direct distribution-based metrics of both computational and experimental results using UQ methodologies and their error ranges. (See also the tasks for process steps 4 and 5).[26] In this step, consideration should be given to the relationship between model fidelity and error levels, and it should be recognized that outputs are stochastic, as opposed to discrete values. Practitioners should review in advance any applicable rules/guidelines for quantifying measurements, e.g., measurement of dimensions, and associated errors pertaining to the experiments at hand. Sources of uncertainty should be reviewed and confirmed to satisfy the validation requirements.

Task 3. Quantify validation metrics and produce data visualizations

In this task, modelers should aggregate the quantitative validation metrics and produce graphs and other data visualizations that illustrate any divergence or disparities between simulation results and the experiments. This can be applied to either discrete data groups and/or to more qualitative trends.

Task 4. Define validation limitations and assumptions

This task again would involve a coordinated effort between some combination of computational modelers, experimentalists, users of the model, and/or computational/simulation analysts. The validation limits imposed by the model itself would be outlined and assumptions about model accuracy or acceptability for applications of interest would be described. In a similar fashion, the experimental validation limits need to be outlined, as well, and any limits of the conclusions and outcomes from the entire validation exercise should be defined, for instance, via a parametric description of the physical regime under examination.

Task 5. Designate funding/resources for developing validation metrics

To engage in a robust validation effort, it is important to ensure that funding/resources are specifically designated to support these efforts, i.e., for developing the validation metrics, which includes metrics that are both closely and poorly matched. The modelers, users of the models, and/or experimentalists involved in the validation effort need to engage customers or regulators to define the thresholds or extent of validation required, in support of approaching funding organizations or other allocators of resources in their organizations. Related to this effort, validation decision support tools can be developed or deployed to help downstream users assess validation results. An example of such a tool is the phenomenon identification ranking table (PIRT), where connections between validation, modeling, and application needs can be mapped and prioritized.

Task 6. Document and disseminate validation metrics among all stakeholders

It is imperative that validation metrics resulting from the validation tasks outlined above be fully documented and disseminated to all parties with an interest or stake in the computational modeling effort. Computational modelers should take the lead responsibility in ensuring that this is done.

III.3. Predictive Capability

It is important to consider the predictive capability of the computational model, particularly as it relates to model validation. Predictive capability is a term used within certain simulation communities to distinguish between model validation, i.e., assessment of model accuracy by comparison with experimental measurements, and model prediction. Predictive capability specifically refers to the capability of the computational model to foretell the response of the system for conditions where experimental measurements are *not* available, or they are *unknown* to the modeler. Predictive capability fundamentally implies a blind prediction, as opposed to replication of previously obtained experimental measurements. The distinction between “Validation” and “Predictive Capability” is depicted in Figure 2 on the next page.

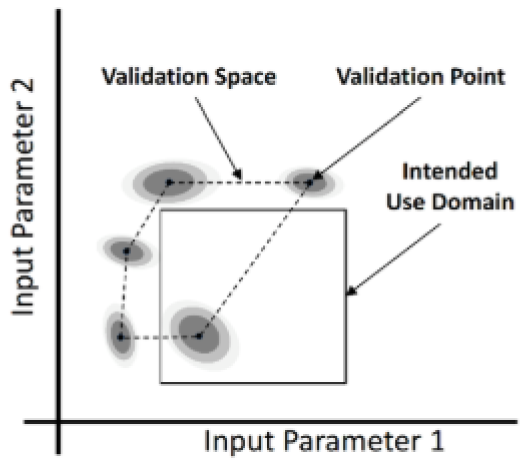


Figure 2: Illustration depicting the difference between predictive capability and validation, via the distinction between the application (intended use) domain, and the validation domain. Figure 2 is taken from an ASME V&V standard.[8]

The validation conditions in Figure 2 are shown as points in the two-dimensional input space of model parameters. The gray region around each point indicates that uncertainty in the input conditions can occur in the experimental conditions, which is then characterized as uncertainty in the input conditions in the simulation of the experiment. The validation domain is defined as the convex hull over the input parameter space where experimental measurements are available. The application (or intended use) domain indicates the region of interest for the real/physical system of interest. These domains are defined in a multidimensional input parameter space, where only two of these dimensions are shown in Figure 2, but in most problems of interest, the dimensionality of the input space is quite large. Model predictions, i.e., blind predictions by definition, typically occur outside of the validation domain because no relevant experimental observations are available. Model predictions also can be computed at other points in the validation domain where experimental data is not available.

Inside the validation domain, a prediction can be viewed as an interpolation within the input parameter space; whereas outside the validation domain, the concept of extrapolation is appropriate. Whether inside or outside the validation domain, the mathematical model of the system of interest is used to predict the response of the system. That is, the accuracy of the interpolation or extrapolation is based on the fidelity of the physics that is represented in the mathematical model, in combination with the statistical features of the input parameters represented in the model. The input parameters for the model are associated with the system of interest and the surroundings to the system of interest. Examples of input parameters associated with the system of interest are geometry characteristics and material properties of the system. Examples of the surroundings to the system of interest are the boundary conditions and excitation conditions for the system. Model input parameters can be either deterministic or stochastic, e.g., parameters characterized by a probability distribution.

One reason that the concept of predictive capability is so important is that, when working with a specific computational model, the practitioner must carefully consider the limits of the predictive capability of the model. The predictive capability of the model should be viewed in terms of the uncertainty sources that contribute to the total uncertainty in the system response quantities of interest. The most important sources of uncertainty are commonly the uncertainty in the input parameters of the model, the uncertainty due to numerical errors incurred in the numerical solution of the mathematical model, and model-form uncertainty. Model-form uncertainty is the uncertainty due to the assumptions and approximations made in the formulation of the mathematical model of the system of interest. These diverse sources of uncertainty contain both aleatoric uncertainty (random variability) and epistemic uncertainty (uncertainty due to lack of knowledge). For a more detailed discussion of predictive capability, see the references section.[2,43,44]

IV. Challenge Problems and Sustainable Funding Programs

Two of the recommendations stemming from the previous TMS workshop and report on verification and validation[3] were associated with (1) developing new challenge/benchmark problems and (2) developing sustainable funding programs for collaborative V&V projects. Both recommendations will help accelerate more robust and widespread V&V practices in MOMS-related communities. Although the need, potential payoff, and general pathways toward development of such challenge problems and sustainable funding programs were discussed previously,[3] the current study teams used this information as a foundation to build recommendations for specific challenge problems and funding programs that might be addressed in the near future by the readers of this study report and/or their colleagues.

IV.1. Challenge Problems

In the initial TMS workshop and final report on V&V,[3] it was suggested that V&V benchmark problems could expose and/or clarify good V&V/UQ practices in MOMS computational modeling and should be limited to simple test problems, as opposed to those of high complexity. In this way, modelers across different research groups can directly compare results, assess the V&V/UQ approaches employed, and learn from each other.[3] Challenges in this context thus refer to efforts that will contribute to benchmarking the efficacy of computational models to predict some set of properties or physical behavior, i.e., they will contain elements of round robin benchmarking challenge competitions with blind predictions (usually including an associated symposium). Developing predictive computational capability for V&V requires scale-aware variational frameworks in which mathematical constructs of the error estimation methods are not only cognizant of the presence of scale effects, but they are also variationally synchronous with the mathematical structure of the problem.[45–47]

Domain-specific challenge problem events have proven successful for community building and teaching essential lessons within and across interdisciplinary teams in numerous technical fields, the most successful of which incorporated significant V&V/UQ efforts. Organizers identify relatively simple test cases and provide guidelines for proper execution of V&V/UQ approaches. The challenge problems should provide a vivid example of existing gaps in the ability to model the behavior of materials and structures. Examples of past benchmark challenges include the 2012, 2014, and 2016 Sandia Fracture Challenges[11–14] and the 2017 benchmark challenge on additive manufacturing (AM-Bench 2018).[48]

Benchmark challenge problems in the MOMS domain will help promote and/or accelerate efforts to implement V&V of computational models within the solid mechanics and materials science communities. In the earlier recommendations to develop such benchmark challenges, some overall action plan guidance included four key tasks or elements: (1) identify key challenges and form interdisciplinary benchmark teams; (2) define the scope and specifics of the benchmark challenge problems; (3) launch the challenge event; (4) select a data repository facility to house the benchmarking information.[3] These generic tasks for any benchmark challenge were discussed[3] in terms of what types of personnel/organizations might lead them, e.g., professional societies and/or national laboratories, a general timeline of how they might be launched, and some implementation suggestions.

The purpose of this section is to expand upon and recommend some specific, potential benchmark challenge problems to be undertaken in order to accelerate implementation of V&V practices in MOMS-related communities. Some of the criteria considered in outlining specific recommended challenge problems included: (1) choosing topics that are relatively well-known, common, and/or previously solved in the MOMS and/or MOMS-related communities; (2) focusing on variabilities and uncertainties in the simulation and experimental results; (3) making sure that V&V/UQ efforts employed are presented by the challenge participants, e.g., making it part of the evaluation/scoring criteria. Some specific elements considered for the challenge problems included possible material types, specimen geometries, process/manufacturing conditions of the components involved, required model input data for the simulations, targeted properties, and model outputs.

For each of the recommended benchmark problems, it is critical that the computational modelers who participate clearly document and/or present all V&V/UQ activities that are employed when they perform their computational modeling. Furthermore, it is also vitally important that the modeling teams clearly explain how they determined input data needed by the model (which were not initially measured during the challenge problem formulation).

If there is interest by a reader of this report in getting involved in organizing any of the benchmark challenge problems, it is recommended to use the overall implementation guidance provided in the earlier report[3], as well, and to use as a previous example or template the types of detailed specifics laid out by the organizing teams of previous benchmark challenges, such as the three Sandia Fracture Challenges[11–14] and/or the 2018 benchmark challenge on additive manufacturing (AM-Bench 2018).[48] As just one example in this regard, the benchmark challenge organizing team would have to determine, in advance, who would make the physical specimens and do the primary characterization/testing experiments for comparison with the computational model results provided by each of the challenge participants.

The challenge problems that were identified are listed below and then elaborated upon:

1. Additive Manufacturing - Variability in Mechanical Properties
2. Hardness and Residual Stress in Joined Parts
3. Crystal Plasticity Modeling
4. Interfacial Friction
5. Precipitation Strengthened Alloys
6. Multi-Step Material Modeling

Beyond these recommendations, readers are encouraged to identify additional challenge problems that will promote incorporating V&V/UQ into computational models associated with the mechanics of a variety of materials and structures throughout structural and soft materials domains.

IV.1.1. Additive Manufacturing - Variability in Mechanical Properties

Additive manufacturing (AM) is an emerging technology with great potential for the fabrication of components with highly complex shapes; however, the wide range of mechanical properties variability in additively manufactured materials has been well-documented.[49–51] Current research efforts in the computational modeling of the complex physical phenomena involved in AM show modeling as a promising path forward for determining and optimizing the mechanical properties of AM components.[52] These methods need a systems approach to material modeling to synchronously connect microstructure and property evolution with the kinematics of the material deposition processes integral to the AM process. The suggested challenge problem would build upon the 2018 benchmark challenge conference on additive manufacturing (AM-Bench), run collaboratively by the National Institute of Standards and Technology (NIST) and TMS.[48] The recommendations provided here could be considered by the organizing committee of the next iteration of that AM-Bench challenge. This challenge could focus on modeling the influence of some number of factors on tensile properties of an AM component. These factors could include: (1) variability with size scale and part orientation; (2) melt pool geometry; (3) fusion line defects and porosities; (4) morphological growth directions; (5) grain shapes and dendritic microstructures. Moreover, the use of specific model parameters and/or surrogate models to inform higher-fidelity models, which achieve greater levels of complexity (through the assistance of high-performance computing solutions), may also be explored. Multiple material classes can be employed, ranging from a variety of metal alloys, e.g., Ti-6Al-4V, 316 Stainless Steel, or 625 Inconel nickel-chromium alloy, to numerous polymers. The choice of specimen geometry should be based primarily on experimental datasets available for validation, with a preference toward tensile samples of multiple sizes with relatively simple part geometries, to minimize multi-scale property gradients and bi-axial or triaxial loading. Depending on the specimen material, the initial processing conditions utilized can be identified, such as the type of AM manufacturing process, e.g., powder bed fusion (PBF), directed energy deposition (DED), or extrusion-based fused deposition modeling (FDM). The goal of the computational models could be to predict performance targets related to a number of specimen properties, such as fatigue, fracture toughness, yield strength, ultimate tensile strength, residual stress/distortion, percent elongation, and/or various microstructural properties defined as gradients or statistical distributions (to appropriately reflect AM part complexities). Such a problem may take the organizing committee considerable time to scope and finalize the details; however, the 2018 AM-Bench challenge and conference[48] provides a good model from which to build.

IV.1.2. Hardness and Residual Stress in Joined Parts

Predictive computational modeling presents great opportunities for accelerating the development of manufacturing innovations related to the joining of materials.[53] It is well-known that development of residual stress, especially during welding procedures in which the temperature of the material(s) is rapidly raised and then cooled during joining of dissimilar materials, is a major issue that needs to be addressed. It can lead to stress-assisted cracking of joints, as well decohesion of the joint interface itself.[54] Since computational modeling of joining presents many complexities and challenges, due in part to steep thermal and compositional gradients, it will be best to focus this benchmark challenge on one specific type of joining process (or two at most) and simple, more easily modeled joint geometries. Gas Metal Arc Welding (GMAW) is a good candidate for the joining process for such a challenge, as it is widely used and studied in the joining of steels and other metals [55–57] and is associated with a long history of computational modeling.[57–59] Perhaps two different welded plate geometries of increasing but not too ambitious complexity could be used in the challenge, e.g., a butt joint and a tee joint. Candidate materials include HSLA-100 steel and/or probably the most common titanium alloy - Ti-6Al-4V. Model inputs would include the welding parameters, e.g., voltage, translation speed, base plate composition, any shielding gas employed, etc., and outputs would include residual strains/stresses, i.e., measured strains could be used to validate the residual stress model. Characterization/testing would include microhardness measurements/mapping, and measurement of residual stress by either neutron diffraction, x-ray diffraction, ultrasonic methods, and/or strain gauges (elastic strain release). Microhardness can be an indirect/empirical indicator of strength and/or the existence of potentially deleterious microstructures, e.g., in steels, a high volume fraction of untempered martensite, which is susceptible to cold cracking.[55,56]

IV.1.3. Crystal Plasticity Modeling

This challenge problem is centered around predictive computational modeling that employs crystal plasticity modeling[60,61] of the stress-strain response and/or microstructure/texture evolution of a single-phase material. Any number of single phase candidate systems can be considered for this challenge, including (1) a β-titanium alloy or low-carbon ferritic steel [body-centered cubic (BCC) grains]; (2) α-titanium [hexagonal close-packed (HCP)]; and/or (3) nickel, aluminum, or an austenitic stainless steel [face-centered cubic (FCC) grains]. Initial processing conditions should be provided for a stress-relieved specimen, which is then subjected to (1) uniaxial tension (~5%); (2) compression; or (3) pure torsion. More than one of these conditions could be modeled and/or tested as different components of this challenge. The specimen geometry(ies) would be dictated by the mechanical test(s), and the specific geometry would be defined by the challenge organizing committee, e.g., dog bone specimens for tensile tests. Characterization would include measurement of the load-displacement curve and characterization of the microstructure (including texture) by techniques such as electron backscatter diffraction (EBSD). The majority of EBSD measurements might be made in two dimensions only, but some serial sectioning in conjunction with the EBSD measurements could provide valuable 3D grain and texture information, as well, for a fuller comparison to 3D crystal plasticity models. Digital image correlation (DIC)[62] also could be used to make surface strain measurements. Goals for the computational modeling predictions would include prediction and direct comparison to the measured load-displacement (or stress-strain) curve, texture/microstructural measurements via EBSD, and surface strains measured by DIC.

As mentioned in the introduction to this section, it is critical for the computational modelers who participate in this challenge to clearly document and/or present any and all V&V/UQ activities that are employed in the development of the computed outputs, e.g., stress-strain response and/or microstructure/texture evolution, that are compared to the experimented measurements.

IV.1.4. Interfacial Friction, i.e., sliding of two surfaces

This challenge problem anticipates predictive computational modeling that employs finite element methods[63] to calculate the interfacial friction between two "simple" materials. The multi-phase candidate systems that can be considered for this challenge are carbon fiber or other polymers, with initial processing conditions performed (1) at room temperature (as heating creates potential instability); (2) under uniaxial or cyclic loading; and (3) while observing interface structure and evolution during curing. The size of the specimens, as well as the number of tests performed, will impact both the time investment and the reliability of the experimental results. With this in mind, the development and mechanical testing of a plethora of very small (micron-scale) samples is recommended. Characterization could include in-situ measurements of microstructural evolution during loading at the two-dimensional interface by way of x-ray diffraction techniques. Moreover, ex-situ serial sectioning of samples could provide valuable 3D grain and texture information, as well, for a further comparison with 3D FE models. More than one of the processing conditions could be modeled and/or tested as different components of this challenge. MOMS-related models could include (1) friction modeling; (2) fracture modeling; (3) contact conditions; (4) bulk material models, and (5) models for contact/slide. Again, it is critical that the computational modelers who participate in this challenge clearly document and/or present any and all V&V/UQ activities employed in the development of the computational outputs.

IV.1.5. Precipitation Strengthened Alloys

This challenge is focused on predictive computational modeling of the solid-state precipitation process of strengthening precipitates and the effect of those precipitates on the mechanical response of the material. Candidate material systems for consideration include (1) an age-hardenable aluminum alloy, such as AA2024 [with Al_2Cu and/or Al_2CuMg precipitates] and (2) an austenitic stainless steel, such as 316L [strengthened by alloy carbide precipitates ($M_{23}C_6$)]. Different annealing/temper treatments can be used, resulting in different precipitate distributions. Microstructural characterization experiments could include some combination of optical microscopy for the matrix grain characterization, electron backscatter diffraction analyses for both the grain and precipitate structures, and transmission electron microscopy (TEM) and/or atom probe tomography for precipitate characterization. Microstructural quantities measured would thus include matrix grain size and precipitate number density, size, and spacing. Examples of possible computational models employed in this challenge are phase field and/or CALPHAD modeling to predict precipitate distributions and any number of precipitation strengthening models (see references[64,65]) to predict tensile strength and/or microhardness (compression). Mechanical testing would thus consist of uniaxial tensile tests to provide yield strength and elongation results and standard microhardness testing to obtain hardness values.

IV.1.6. V&V for Multi-Step Material Modeling

In mechanics and materials communities simulations often span vastly different length scales.[18] While this issue has proven to be quite challenging over the years, recent advances in machine learning have greatly enhanced scale bridging for predicting microstructure and property evolution in materials, by integrating from density functional theory through statistical mechanics to continuum phase field modeling.[66,67] This challenge problem is designed to explore predictive computational modeling that employs the use of multi-step material modeling approaches, integrating two models, with the output of one model serving as the input of the next model. While this challenge problem is designed to highlight the importance of uncertainty quantification, one or both of the models also must have well-established V&V underpinning. An interesting material for this challenge would be a blade or disk of some type of Ni-based superalloy which was processed via (1) casting; (2) forging; or (3) additive manufacturing. The latter processing method would result in a multi-step material modeling problem that entails an intricate link between material evolution and the kinematics of the printing process, that has a direct impact on the properties of the resulting material and the long-term fatigue life of the product. The ultimate goal of this exercise would be to model a macro-scale material property, such as fatigue, rupture, or creep; however, to do so, participants might be required to first utilize an "intermediate" model to predict the expected microstructure of the materials and employ that resulting microstructure to simulate the behavior of the final material property. Based on available experimental data, the initial processing conditions will be determined by the challenge organizing committee and subsequently provided to the challenge participants. More than one of the processing conditions could be modeled and/or tested as different components of this challenge. It should be recognized that any success will be specific to the scales and phenomena that are being brought into play. Furthermore, it is especially critical that the participating computational modelers clearly document and present all V&V and UQ methodologies and activities utilized in the development of both the computational models and resulting outputs. Any viable teams will likely require the active collaborations of modelers, experimentalists, computational analysists, etc. (see Section III.2).

IV.2. Sustainable Funding Programs

IV.2.1. Some Current and Past Efforts

To incorporate V&V/UQ practices into established scientific workflows and sponsor widely available and impactful challenge problems, it is necessary to obtain sustainable funding for these large-scale, collaborative efforts. The successes and failures of other communities that have already seen the need to invest in V&V/UQ for their scientific enterprises can provide a good template for new efforts. Thus, some past/current efforts in this regard are now considered.

As one example, regulatory communities have invested heavily in V&V to support simulation and modeling efforts within their domains. The DOE's National Nuclear Security Administration (NNSA) Advanced Simulation and Computing (ASC) program has developed and deployed computational simulations to analyze and predict the performance, safety, and reliability of nuclear weapons since 1995, with V&V remaining an essential component of that program to this day.

This program, originally driven by the 1992 ban on underground testing of nuclear material, employs computational modeling and simulation as the dominant contributor for providing design confidence along with limited non-nuclear experimentation and testing. In this case, the push to implement vigorous V&V was driven by the great and exclusive need driving the sustained funding.

Among the plethora of existing computing and simulation-centric research and/or development programs, most do not specifically call out V&V/UQ in their program descriptions, yet several are engaged in activities that would benefit by incorporating V&V. For example, the National Strategic Computing Initiative (NSCI), whose mission is to "accelerate scientific discovery and economic competitiveness by maximizing the benefits of high-performance computing (HPC) research, development, and deployment," does not explicitly reference V&V/UQ; however, the implementation of V&V/UQ techniques and practices are vital to truly maximizing the benefits of HPC. Consequently, this is a large-scale program which offers a structure for proof of V&V/UQ concepts, even though it does not initially purport to do so. In this case, not incorporating V&V/UQ into the scope/workflow of such projects could jeopardize the integrity of the scientific and economic investment that this and other similar initiatives represent. In bringing the full power of computational modeling to bear on programs and products, V&V/UQ is an essential element for confidence in decision making and fully realized value.

IV.2.2. Recommendation for New Programs

Within MOMS-related communities, the increasing investment of modeling and simulation development can be jeopardized by the lack of uniform qualification metrics that are provided by rigorous V&V methods. As discussed in Section II, the practice of V&V/UQ allows for a systematic way of achieving and conveying the confidence needed to utilize these models and simulations to their fullest potential. Provided below are recommendations for just a few specific program areas in which funding organizations would be well-served by supporting independent V&V/UQ research programs, and/or including V&V/UQ practices within their current activities, to ensure effectiveness in funding and confidence in the results. These recommended program areas, however, are by no means all-inclusive, as many other currently funded arenas could also benefit from such V&V/UQ activities.

National Defense-Related Challenge Problems

As discussed in Section IV.1 and observed with the Sandia Fracture Challenge, discipline-wide and interdisciplinary challenge problems are powerful ways of demonstrating the impact of V&V/UQ. As seen in multiple past challenges[11–14], while specific V&V/UQ practices are rarely required, modelers who employed these practices performed best in the challenges. This is especially true for challenge problems that require the use of nondeterministic models and simulations which produce different outputs in different runs with the same inputs. These simulations are used in predictions to find answers to questions in unknown spaces, where the confidence bestowed by rigorous V&V/UQ provides great benefits. It is precisely this ability to quickly and reliably produce answers to challenging problems that makes these simulations so appealing; however, the answers simply cannot be consistently relied upon without the implementation of rigorous V&V/UQ.

Many challenging materials problems are of particular interest to the Department of Defense (DoD), since material properties' responses (such as fatigue, corrosion, stress cracking failure) to outside stimuli are of critical importance in military applications. Today, remaining on the cutting edge of military equipment and technology requires accelerating the discovery-testing-prototyping process to "deliver performance at the speed of relevance"[68] to the field. Nondeterministic modeling would significantly lessen the need for repetitive testing; however, accuracy and quality of the modeling is paramount. V&V/UQ is the only way to ensure robust confidence in such nondeterministic testing. Underlying this, therefore, is the requirement that V&V/UQ methodology be rigorous and consistently replicable to maintain simulation confidence, in support of America's militarily competitive edge. The great need for successful solutions should highlight the importance of the implementation of V&V/UQ techniques to any DoD-sponsored computing effort, as well as drive commensurate, elevated funding investments in this area. Such funding would then support the sorely needed acceleration of the broad implementation of robust V&V/UQ practices in the MOMS arena.

Materials Genome Initiative

Since its inception in 2011, the multi-agency Materials Genome Initiative (MGI) has supported numerous efforts to discover, manufacture, and deploy advanced materials twice as fast and at a fraction of the cost by combining computation, experiments, and digital data.[5,6] The merger of these three components as part of a materials innovation infrastructure makes the MGI process ripe for supporting V&V/UQ, as it is imperative that the resulting outputs are reliable and replicable in order to successfully achieve the desired acceleration and cost reduction in advanced materials development. The implementation of V&V dovetails well with, and is critical for, the success of this MGI-based, computational materials design approach, as it naturally builds confidence when simulating previously unexplored space through the integration of computational and experimental data. Therefore, agencies with significant investment in the MGI should be incentivized to support independent and/or supplemental programs designed to strengthen V&V/UQ, particularly in relation to existing projects.

Technology Transfer

Moving technology from laboratories to the marketplace is a challenge following discovery and is vital to our continued technological and economic growth. As such, the federal government has multiple, ongoing, national initiatives to help facilitate this transition, including the Department of Energy (DOE)'s Office of Technology Transitions, Small Business Innovation Research (SBIR), and Small Business Technology Transfer (STTR) programs, the National Science Foundation's (NSF's) I-Corp program, and the Federal Laboratory Consortium for Technology Transfer (FLC). While many of these programs focus on providing infrastructure, entrepreneurial advice, and access to resources and networking opportunities, technology transfer models oriented toward improving the planning and execution of tech transfer programs also have been considered.[69,70] The implementation of V&V in the MOMS domain has the potential to propel technology transfer on multiple fronts. First, the use of verified and validated MOMS-based research simulations would accelerate the rate of scientific discovery (as discussed in Section II) supporting this technology transfer. Secondly, as specific V&V/UQ methodologies are developed, technology transfer vehicles would serve as yet another tool to effectively and efficiently translate scientific innovation in trusted computational models into economic success.

Any agency looking to simultaneously promote rapid and reliable laboratory-scale innovation and industry-scale application would benefit greatly from investing resources for (1) V&V/UQ implementation within existing research funding and (2) novel research into V&V methodologies specifically geared toward predictive computational modeling supporting technology transfer.

Artificial Intelligence

Artificial intelligence (AI) represents a new frontier of technological advancement which is seen as crucial for the worldwide science and engineering enterprise and in nearly all sectors within the 21st-century economy. The prospect of leveraging computing power to extract meaningful insights from the sea of available data is immense. As a result, researchers, companies, and nations are investing untold amounts of time and money to rapidly develop the latest algorithms which will best-augment human understanding. However, the exceptional promise of AI comes with exceptional risks.

At their best, algorithms in machine learning, the sub-discipline of AI that is most widely in use in engineering, ingest massive amounts of raw data and then use it to predict some aspect of future events, without the need of governing equations or human intervention. However, the more reliant we become on the "black box" oriented solutions of machine learning, the more vital it will be to have some external method of confirming that such predictions are yielding credible results. Thus, strong V&V/UQ practices are needed to support and enable the widespread practical applicability of machine learning. V&V thus offers a path for machine learning to become more scientifically grounded. However, as we move from a mathematical, methods-driven approach of predictive modeling to a data-driven, machine learning approach, it is neither trivial nor obvious how current V&V/UQ practices will be fully adopted in this arena.

The development of V&V/UQ methods tailored specifically for AI is an open research question that merits dedicated research funds and efforts. As future, major initiatives are likely to focus on AI, it is, therefore, critically important for funding agencies to establish and support independent research programs that dedicate significant resources to the research and development of novel AI V&V/UQ methodologies. At the very least, this activity should be viewed as an equivalent priority with all other AI-related funding efforts, as it represents an essential component of achieving long-term AI success. Moreover, in this rapidly developing field, the connection of V&V/UQ practices to AI applications could lead to the development of new tools for widespread use across the greater AI community. In an interdisciplinary sense, therefore, investment in V&V/UQ for AI would have a positive effect not only for materials science and mechanics-related communities, but for a number of emerging, nearly ubiquitous AI-related communities as well.

V. Symposia, Conferences, Workshops, and Studies

Domain-specific symposia, conferences, workshops, and studies could strongly contribute to accelerating the adoption of more robust V&V/UQ practices in the MOMS community. Some recommended offerings for stimulating such V&V/UQ adoption and implementation are provided below. The topics associated with these offerings could apply to either specific elements of V&V methodologies themselves or specific technical domains of the MOMS computation models. An example of the former would be a workshop or symposium/session on solution verification at a professional society conference and an example of the latter would be a study, workshop, and/or symposium on V&V and UQ practices associated with the computational modeling of metal fatigue.

V.1. Symposia and/or Conferences

Although some organizations already run V&V/UQ symposia, it is recommended that more professional societies and other organizations develop symposia to stimulate interest and foster knowledge sharing, adoption, and implementation of V&V/UQ practices, specifically in the MOMS community. These organizations could work with appropriate volunteer organizing committees to develop and implement such symposia. Depending on the interest level and projected number of attendees, symposia could be centered around, not only V&V/UQ technical subdomains of the computational modeling of material and structures, but on topics related to specific elements of V&V/UQ methodologies themselves, as discussed in section III of this report. Possible methodology-specific symposia topics are UQ in simulations and validation experiments for MOMS computational models. Nevertheless, in all cases, symposia organizing committees would dictate the specific symposium topic and scope. They could use the concepts and recommendations in this report in conjunction with ideas that resonate best with their research communities and technical committees, to decide what would be most needed, impactful, and well-attended.

Workshops (see the subsection below) also may be held in coordination with the symposia, at the same event where the symposia are executed.

Some professional societies and other organizations that either already run V&V/UQ symposia or that can initiate new symposia related to MOMS computational modeling will now be considered.

The **American Society of Mechanical Engineers (ASME)** has a strong, rich history in V&V efforts and committees and already runs a popular symposium on V&V (see https://event.asme.org/V-V). Other ASME conference venues that might include symposia related to V&V of computational models of materials and/or structures are the Smart Materials, Adaptive Structures and Intelligent Systems (SMASIS) conference, as well as the ASME Annual meeting.

The Minerals, Metals & Materials Society (TMS) could work with appropriate volunteer technical committees to develop V&V symposia in the MOMS arena, as well. Such TMS committees might include but not be limited to: Integrated Computational Materials Engineering (ICME); Computational Materials Science and Engineering; Advanced Characterization, Testing and Simulation; Process Technology and Modeling; Mechanical Behavior of Materials; and Shaping and Forming. Some TMS events where V&V-related symposia would be most appropriate and impactful include the TMS Annual Meeting & Exhibition, the Materials Science and Technology (MS&T) Technical Meeting and Exhibition, and the World Congress on ICME. TMS also could collaborate with the organizers of the World Congress on Computational Mechanics (WCCM) to hold a session/symposium related to V&V of models of materials and structures at that event.

The **American Institute of Aeronautics and Astronautics (AIAA)** has also been very active in holding short courses/workshops/symposia on V&V and/or ICME and could hold future symposia centered around V&V/UQ activities associated with computational models for aerospace materials and/or aerospace structures, perhaps at their AIAA SciTech Forum and Exposition.

The **Engineering Mechanics Institute (EMI)** holds the EMI Conference and Probabilistic Mechanics & Reliability Conference, which is an excellent venue for V&V symposia in the solid mechanics realm. The **Society for Industrial and Applied Mathematics (SIAM)** Conferences on Mathematics of Data Science (MDS) and Computational Science and Engineering (CSE) also could be good venues for such targeted V&V symposia. The **U.S. Association for Computational Mechanics (USACM)** has a Technical Thrust Area on Uncertainty Quantification and Probabilistic Modeling, which could lead V&V symposia at the US National Congress on Computational Mechanics (USNCCM). The **International Conference on Multiscale Materials Modeling (MMM)**, which is in its 10th iteration and is typically planned by organizing committees consisting of volunteers mostly from academia, is also an excellent candidate to hold symposia/sessions specifically targeted toward V&V and UQ practices in computational models associated with the mechanics of materials. The **International Conference on Plasticity, Damage & Fracture** and the **International Symposium on Plasticity and Impact Mechanics (IMPLAST)** also are venues ripe for sessions focused on such V&V practices.

Many government regulatory agencies also have strong efforts in support of computational modeling and, thus, could organize V&V symposia. The **Federal Aviation Administration (FAA)** already holds an annual Verification and Validation Summit, geared toward FAA or related government agency employees or contractors. The **Food & Drug Administration (FDA)** also has strong computational efforts and might either lead or coordinate with another organization on V&V symposia, centered about computational modeling of "soft materials." The **Nuclear Regulatory Commission (NRC)** actively addresses licensing issues for new reactors and is attuned to the importance of and opportunities for V&V in nuclear regulatory processes. They organize various conferences and symposia (see https://www.nrc.gov/public-involve/conferences.html) and thus might support V&V symposia in the MOMS realm.

Other government agencies are involved in V&V-related conferences and workshops and might have interest in supporting such V&V symposia. For example, a future iteration of AM Bench,[48] a benchmark challenge conference on additive manufacturing run collaboratively by the **National Institute of Standards and Technology (NIST)** and **TMS,** could include a session specifically geared toward presentations on V&V and UQ practices related to the computational modeling performed as part of the challenge. **Sandia National Laboratories** has a long history of leading V&V/UQ efforts[15,16], including their Verification and Validation (V&V) program (www.sandia.gov/ASC/verification_validation.html) and the Sandia Fracture Challenge.[11–14] They could be instrumental in planning such symposia, perhaps in association with the next iteration of the Sandia Fracture Challenge.

The **National Science Foundation (NSF)** and other federal funding agencies, such as the **Air Force Office of Scientific Research (AFOSR),** the **Office of Naval Research (ONR),** the **Department of Energy (DOE),** the **National Institute of Health (NIH)**, the **Army Research Office (ARO),** and the **Defense Advanced Research Projects Agency (DARPA)** support various workshops, conferences, and/or symposia on a wide variety of topics, sometimes related to V&V. Thus, they could be instrumental in providing some support to assist with travel and registration for students, young investigators, and/or other speakers who might not otherwise be able to attend the types of V&V symposia suggested above.

V.2. Studies/Workshops

Some recommendations are now considered to further advance acceleration of robust V&V adoption within the MOMS community by developing highly impactful, robust science and technology accelerator studies, as well as by developing smaller/single workshops.

Future Studies

These targeted, robust studies should be output-oriented activities with tangible deliverables; for prior examples of such studies, see past TMS science and technology accelerator study reports.[18,21,53,71–73] Depending on the specific topic, it might be valuable for these initiatives to be arranged by cross-organizational, interdisciplinary teams. Although the individual study initiatives might have domain-specific target objectives, they also could be designed to address multidisciplinary aspects of V&V/UQ for the MOMS community.

These studies should have specific topics and target objectives that would be highly impactful in stimulating and/or supporting the widespread implementation of verification and validation in computational models of the mechanics of materials and/or structures. In this regard, **some recommended study topics/target objectives are**:

- Design and implementation of validation experiments for MOMS-related computational modeling
- Building awareness with industry managers in the MOMS community whose organizations/ teams may stand to benefit from newly implemented V&V/UQ approaches
- Educating/training V&V/UQ community members across industry, national laboratories, and the broad research community in new skill sets and tools that are specifically relevant to current technological progress in materials and solid mechanics
- Studies that motivate action around highly detailed, community-developed V&V/UQ strategies, methods, and best practices, i.e., around specific elements or process steps of the recommended practices framework provided in section III.

Selection of the organizations and/or personnel who should be involved in both leading and participating in such studies is critical. Individual participants should include active researchers and practitioners with computational modeling and experimentation backgrounds from all organization types, including academia, industry, and government research laboratories. Other stakeholders with interests in this arena, such as industry managers in the MOMS community whose buy-in is required to commit the time and resources needed in implementing these V&V/UQ efforts, also should be involved. Organizationally, professional societies would be appropriate for convening the requisite experts and organizing and guiding these workshops and studies.

Workshops

Smaller, single-day workshops are also a good vehicle for advancing V&V/UQ in the MOMS community. These would most likely be organized by professional societies, perhaps sometimes in collaboration/conjunction with the symposia mentioned earlier. Thus, these workshops could be held in conjunction with various conferences that would be appropriate venues for hosting working sessions on specific V&V/UQ topics. Some examples include TMS's Annual Meeting and Exhibition, the World Congress on ICME, ASME's annual V&V symposium, AIAA's SciTech Forum and Exposition, and/or SIAMs conference on the mathematics of data science. In addition to including researchers and V&V/UQ practitioners with computational modeling and experimentation backgrounds from academia, industry, and government research laboratories, such workshops also should involve software developers (both commercial and within national laboratories), as they can play an important role in training V&V users on specific V&V/UQ topics.

Some recommendations for specific workshop topics include:

- Connecting UQ across multiscale simulations and materials characterization experiments
- Sources and implications of experimental UQ on V&V of single and multiscale models of materials behavior
- V&V in computational modeling of hydrogen effects/embrittlement in steels
- Workshops specifically focused on the challenges in computational modeling of soft materials, active materials, and biomaterials
- Fusion of data into models - how the data affects the predictive capability of models based in mathematical methods
- Coupling validation experiments and computational modeling, for example:
    - Coupling in situ X-ray diffraction with crystal plasticity modeling
    - Coupling dislocation-interface reaction models with associated in-situ transmission electron microscopy (TEM) and electron backscatter diffraction (EBSD) experiments at different scales of resolution
- Two-level hierarchical and/or concurrent models for microstructure and property evolution, based on multi-modal, multi-resolution experiments
- Models for microstructure and property evolution in Titanium 6AL-4V or 316L stainless steel alloys as a function of additive manufacturing process path and process parameters (for a target notched specimen configuration)
- The principles and challenges of code verification for researchers and practitioners
- Tradeoffs of model complexity and model reliability
    - V&V/UQ experiment-model fusion for overarching systems-level decision support, e.g., in additive manufacturing

VI. Strategies to Connect Modelers and Experimentalists

Challenges Involved

Before proposing a list of opportunities and tactics to achieve the goal of promoting multidisciplinary V&V teams, it is important to identify known challenges that prevent effective collaboration of modelers and experimentalists on V&V/UQ efforts. The following were found to be significant impediments to effective communication between modelers and experimentalists regarding V&V/UQ practices.

One significant challenge revolves around the lack of dedicated funding for dual V&V/UQ workstreams, i.e., experimentalist-modeler teams. In a scientific enterprise which increasingly merges computational and experimental techniques, it is imperative that funds be made available to help promote and formalize the communication between modelers and experimentalists. Moreover, when funds are made available, they typically are dispersed by a single entity, but sometimes also by cooperating funding sources. Projects funded from a single source often focus disproportionally on either the modeling or experimental activities with mere tacit involvement from the other party, instead of cultivating a project of true interest to both groups of collaborators. In the case of projects funded from multiple sources, there are sometimes imperfectly matched goals between the shared funding sources, which can result in projects with "collaborators" who work independently on a subset of goals that may be of much more interest to just one of the funders, and the principal investigators (PIs) often only interact with their fellow investigators to communicate results. This can serve to perpetuate the separation between modelers and experimentalists. In addition, the extent to which different funding sources have different goals can introduce difficulties, such as funding gaps or discrepancies with the strategic outcomes of the V&V efforts in the funded projects.

Other barriers to effective collaboration revolve around team dynamics; such barriers include (1) the inability to geographically co-locate teams, i.e., experimentalists and modelers are typically not located in the same physical location, which if so would be conducive to conducting robust collaborative V&V/UQ work; (2) the misalignment of experimentalist-modeler workflows and timelines, i.e., experimentalists and modelers have fundamentally different constructs in terms of their priorities, disciplines, and languages used, and models tend to progress at much different rates than experiments do; and (3) an aversion to validation results, stemming from the fact that the validation process tends to find shortcomings in both the experiments and simulation codes. For these reasons, many modelers and experimentalists are often unenthusiastic, unable, or reluctant to devote the additional time and resources normally required to collaborate on a robust V&V/UQ project.

Finally, another significant hurdle is a lack of V&V buy-in among industry managers, since many industry managers do not fully recognize the transformational impact V&V/UQ would have on applicability of the modeling and simulation results and, therefore, are not motivated to dedicate specific funding for V&V/UQ efforts.

Recommended Strategies and Opportunities

Considering these and other challenges, a list of opportunities and strategies to achieve the stated goal of bringing together multidisciplinary V&V teams of modelers and experimentalists was developed. Three overarching strategies are identified in the left column of Table 3 and several opportunities, or tactics, are proposed for each of these strategies, as displayed in the right column of Table 3. It should be noted that although the opportunities/tactics in Table 3 are sorted into the three proposed strategy areas, some opportunities may naturally align with more than one of the three strategies.

Table 3: Strategies and opportunities to bring together multidisciplinary teams of modelers and experimentalists

Strategies	Opportunities/Tactics
Provide Incentive Mechanisms that Motivate V&V Contributions	• Multidisciplinary V&V funding streams and/or requirements • Support for publication venues that reinforce joint experimental-computational V&V efforts • Buy-in of collaborative V&V among industry managers • Robust experimentation requirements for challenge problems
Build Awareness of the Value of V&V/UQ	• Value-driven collaborative V&V/UQ partnerships • Professional society V&V committees/working groups • Recognition and achievement awards for V&V/UQ efforts • Advertisements for shared V&V/UQ resources
Reduce the Entry Barrier for New V&V/UQ Users	• Networking events for V&V/UQ team matchmaking • Joint educational opportunities for experimentalist-modeler teams

Strategy #1: Provide Incentive Mechanisms that Motivate V&V Contributions

Multidisciplinary V&V funding streams and/or requirements

An important fundamental change needed is the direction of monetary incentives, no matter the industry or organization. Integral to the long-term success of any project is adequate funding to support it, making such fiscal support a powerful change agent. With that said, the need for dedicated V&V/UQ funding within the MOMS community cannot be overstated and can be best engaged via two complementary methods. First, there is a need for funding of independent, standalone research programs devoted to advancing the robustness of V&V/UQ techniques relevant to the field and promoting the accessibility of experimentalists and computational modelers to these techniques and to each other. This can be achieved by establishing deliverables, goals, and success criteria for both experimental and computational V&V/UQ workstreams, to set outcomes that drive toward transformative impacts for both modeling science and experimental workstreams. Second, current grant programs should include some earmarked funding for V&V/UQ activities, thereby setting requirements or guidelines for funded V&V/UQ work to provide a more holistic, evidence-based view to understanding new phenomena and their mechanisms. Understanding new phenomena explicitly requires V&V/UQ; experimental evidence of the phenomena and a validated model are two synergistic tools for understanding physical behavior, and that understanding is not complete without a validated model. By incentivizing and motivating efforts to conduct robust V&V/UQ work, i.e., deep-dives - not just surface-level analysis, funding requirements could mandate primary or additional tests that are essential to understanding the studied phenomena and that, by their nature, require substantive collaboration between modelers and experimentalists.

Publication venues that reinforce joint experimental-computational V&V efforts

Publications are a key metric for advancement in academic and government laboratories. However, there are currently limited venues and professional benefits to publishing V&V/UQ evidence packages through conventional publication channels. In addition, the amount of time and labor needed to produce such results make the desire to do so even more prohibitive. An increased focus on publication venues that truly/exclusively support joint experimental-computational work is needed. Some existing publication venues, e.g., the TMS journals *JOM* and *IMMI*, issue thematic groups of articles (series), and such thematic sections in journals could be centered around coordination of experimental methods and modeling for V&V, but as of yet essentially no venues take a comprehensive view of V&V/UQ. Publication submissions with strong modeling and experimental components are typically too long for most journal publications. Therefore, journals could produce, for instance, a multi-part publication series that captures comprehensive aspects of modeling and experimentation. Finally, new publication venues could focus on how specific, collaborative V&V/UQ efforts are directly benefiting the broader community by improving widely established (or "standard") experimentation and modeling approaches.

Buy-in of collaborative V&V among industry managers

Collaborative V&V/UQ buy-in among industry management professionals and other decision-makers should be promoted, perhaps by using challenge problems as a vehicle. In this way, industry managers can be shown examples of the near- and long-term benefits of collaborative V&V/UQ efforts and demonstrate how the V&V/UQ process can accelerate certification/qualification processes and improve decision making. In particular, challenge problems might be useful vehicles for enumerating these benefits if oriented toward demonstrating the necessity and benefits of collaboration, especially between modelers and experimentalists. Professional societies and/or government organizations could play a role in organizing challenge problem workshops geared toward industry managers.

Robust experimentation requirements for challenge problems

Stringent experimentation requirements for addressing V&V/UQ-related challenge problems could be set, particularly by considering requirements for challenge problems that mandate additional testing to ensure V&V efforts that produce more than just one problem solution. Imposing rating/judging penalties for incomplete testing would further incentivize the stated goals. V&V/UQ teams also should be given physical samples to experimentally extract data beyond what is initially provided by the challenge organizers. Best or recommended practices for practitioners, like those provided in Section III of this report, also could help encourage generation of robust experimental datasets that benefit the broader V&V/UQ community and that require intimate collaboration between modelers and experimentalists.

Strategy #2: Build Awareness of the Value of V&V/UQ

Value-driven collaborative V&V/UQ partnerships for decision making

The link between value systems of V&V/UQ collaborators, e.g., both individual and corporate, and the impact on the quality of research and/or engineering decisions needs to be brought to light. For example, industry requires reliable information for decision support, while this concept is rarely (if at all) taught at universities. V&V/UQ practitioners largely require realistic metrics to measure the value of their contributions and would benefit from demonstrated examples of how validated or proven models can reduce the need for expensive experiments, improve experimental diagnostics and facilities, and reduce overall development times, budgets, and risk. Funders should highlight the value of individual contributions, as well as collaborative efforts, toward the positive impact on decision credibility. Individuals may be more motivated to participate in collaborative V&V/UQ activities if they believe their contributions have a meaningful impact on the quality of decision making. Thus, creating a mechanism to produce and publish results of cooperative experimentation and modeling efforts in which state-of-the-art V&V/UQ approaches to improve decision support quality are employed would go a long way toward incentivizing experimental-modeling collaborations. Additionally, from the corporation perspective, publication of V&V/UQ evidence packages would ensure a traceable way to have confidence in decisions and to transfer projects and knowledge across divisions and in the face of key staff turnover.

Professional society V&V committees/working groups

Professional societies beyond ASME, ASCE, and AIAA should be encouraged to form V&V working groups or committees to advance, for instance in the MS&E community, high priority ICME and MGI topics. Professional societies are well-positioned to frame V&V/UQ activities around critical foundational problems in topic areas such as ICME and the MGI, where V&V/UQ could deliver a substantial return on investment. Widespread V&V/UQ across the MOMS community will require sustained guidance and involvement from dedicated V&V champions. Professional societies could form dedicated V&V/UQ working groups (within or across multiple professional societies) to demonstrate the benefits of collaborative V&V.

Recognition and achievement awards for V&V/UQ efforts

Recognition and/or prizes could be offered for collaborative work that advances the awareness and importance of V&V for multiple end users. The candidate pool could include, for instance, V&V/UQ practitioners within or across industry, government laboratories, or universities. Recognition of effort could be based on: (1) the value delivered to the user and/or (2) progress made to advance the field, in terms of community-wide V&V goals. As one example, a "young investigator" award for exemplary V&V contributions could be created by funding agencies and/or professional societies.

Advertisements for shared V&V/UQ resources

Advertisement space in journals and/or on websites could be utilized to notify key V&V/UQ practitioners of available/existing resources, e.g., experimental databases and open source tools, and thus stimulate the requisite collaborations between modelers and experimentalists.

Strategy #3: Reduce the Entry Barrier for New V&V Practitioners

Networking events for V&V/UQ team matchmaking

Conference events can be used to encourage cross-pollination of prospective V&V/UQ teams via conference organizers who could host expert panel discussions or sessions designed to help prospective V&V/UQ practitioners forge partnerships between teams of experimentalists and modelers. These networking events also provide the avenue for informal professional mentoring opportunities to shepherd novice practitioners.

Joint educational opportunities for experimentalist modeler teams

Design educational opportunities that allow teams of two or more graduate students to participate in the full V&V/UQ process. Fellowships or scholarships that allot funding for both an experimentalist and a modeler would give students experience in collaborating on V&V teams. Ultimately, one or more "joint" PhD theses between an experimentalist and a modeler may help encourage a culture of collaborative V&V in the MOMS community.

VII.

Strategies for Improving Multidisciplinary Training and Curricula Development

A wide variety of baseline/essential skill sets are needed for conducting proper V&V/UQ in computational models associated with the mechanics of materials and structures. Eight key competency areas were identified, wherein such skill sets are needed. These competency areas include:

- Code Verification
- Solution Verification
- Model Calibration and Calibration Experiments
- Uncertainty Characterization
- Uncertainty Propagation
- Model Validation Experiments
- Validation Metrics
- Predictive Capability

It should be noted that these competency areas map very closely with the process steps (numbered boxes) in the recommended practices framework depicted in Figure 1 of Section III.

<u>Seven mechanisms, or strategies</u>, are recommended for building such skill sets:

1. Integration of V&V/UQ modules into existing core university courses
2. Creation of new core and/or technical elective courses
3. Stand-alone tutorials, short courses, and/or workshops
4. Open source tools development and dissemination
5. Challenge problems
6. Tutorials and texts
7. Mentoring

Each of these recommended strategies will now be considered in more depth.

1. Integration of V&V/UQ modules into existing core university courses

Single- or multi-day lectures or modules could be inserted into existing course offerings. V&V/UQ elements could be added to existing assignments in those courses.

Undergraduate courses: For undergraduate core courses that involve significant components of computer code development, such modules should include preliminary exposure to code verification concepts. Preliminary exposure to solution verification concepts (see section III) could be provided in undergraduate courses that involve learning and application of different computational models to specific engineering problems. Such verification steps could be added to existing assignments in these courses; for example, a verification step could be added to a final project that uses a computational model. Capstone materials and/or structural design courses, for example, may be good candidates for such additions. Preliminary exposure to uncertainty quantification could be provided in undergraduate courses that contain elements that dovetail well with error or uncertainty analyses. For example, a module or lecture(s) on uncertainty quantification could be provided within a core thermodynamics course already offered within an engineering or physical sciences department and/or in basic courses in materials characterization. Exposure to probability and statistics topics could be provided in existing undergraduate laboratory/experimentation courses; this would promote skill set development in the competency areas of model calibration and validation metrics.

Graduate courses: In essentially all existing specialty courses at the graduate level that involve modeling and simulation, VV and UQ concepts could be introduced. Some examples include graduate courses in materials science and engineering and/or mechanical engineering departments that include elements of alloy design, mechanical behavior, integrated computational materials engineering (ICME), or structural design. At the graduate level, such modules/lectures could cover any or all of the seven competency areas listed herein, and the recommended practices section (section III) of this document could serve as a good resource from which to build the foundation for and integration of more advanced elements in these graduate-level lectures/modules. This content would be developed by experts, for integration by non-experts into existing courses.

2. Creation of new core and/or technical elective courses

Inserting new courses into what are already densely packed curricula in most science and engineering departments can be a daunting task for the core curriculum. However, for technical elective courses at the undergraduate and graduate levels, initiating new courses focused on V&V/UQ is recommended. Courses in UQ for multiscale modeling that also deal with V&V principles and methods might be broadly appealing.

Undergraduate courses: Such standalone courses at the undergraduate level do exist in some leading programs today but are extremely rare. Introductory exposure to V&V can be provided, though, even at the undergraduate level, in new dedicated technical elective courses, and it is recommended that more of these types of courses be developed and offered. These could especially be valuable for undergraduates to gain some level of knowledge and proficiency in all aspects of V&V/UQ. Without question, V&V should be a component of instruction in any design or capstone design course for undergraduates, to reflect its value in the workplace.

Graduate courses: It is strongly recommended that dedicated core and/or elective courses in V&V/UQ be offered in many more MOMS-related departments; these courses could address the full spectrum of all seven competency areas mentioned. This is especially important because graduate school is the breeding ground for the next generation of researchers in industry and at government laboratories. In depth knowledge of all aspects of V&V/UQ is absolutely essential to mitigate risk and maximize the great potential benefits of modeling (see section II). Moreover, graduate schools will produce the next generation of university faculty who will not only be developing and using such computational models but will be teaching future generations of practitioners.

Continuing Education: Scientists and engineers already in the workforce also could benefit from new courses via exposure to dedicated graduate- and post-graduate-level education. For example, some universities offer courses in which professionals in the existing workforce can attend for several weeks and earn at least one academic credit.

3. Stand-alone tutorials, short courses, and/or workshops

Any tutorials or short courses for professional development that are geared toward practitioners who are already in the existing workforce are considered part of continuing education. They could run anywhere from two days to two weeks and with anywhere from 10 to 100 individuals per offering. Professional societies have a strong history of building such courses or workshops, which can be taught at existing conference venues, standalone venues, on-site within a company or national laboratory, or in a university setting. In the case of on-site courses taught at individual companies, non-disclosure agreements might be required. Efforts should be made to keep registration costs affordable to promote attendance. Some universities also are encouraging agencies to fund "educate the educator" courses, e.g., summer courses, to help prepare university faculty instructors and key leaders at national laboratories.[72] Corporate sponsorship also could be provided for specific short courses in high demand within various industrial companies.

In-depth domain or sector specific short courses can be offered to support competency areas, including code and solution verification, model calibration, model validation experiments, and validation metrics. Such courses can be offered to individual companies, e.g., on-site, or to attendees from a broad group of companies and other organizations that have interest in these domain-specific short courses (at more neutral venues). The domain-specific areas would be dictated by demand, and such demand could be elucidated via surveys of the community by professional societies. Examples of such domain-specific courses might include courses on the design and application of validation experiments, validation metrics, code verification, and solution verification.

Courses aimed at increasing competency in non-deterministic simulation and uncertainty characterization are more broadly applicable across the entire spectrum of computational models associated with the mechanics of materials and structures (MOMS). These can be provided as short courses on the order of 2-5 days that are delivered by external subject matter experts via professional societies and/or via individual consultancies. These courses would offer broad introductory exposure to key UQ concepts and tools, as well as provide references and resources for practitioners to take back to their organizations, enabling them to dive more deeply into implementation of specific UQ methodologies. It is important to note that university involvement in the education aims of this V&V report must fully address uncertainty quantification to effectively educate and to draw the scientific and engineering interests of faculty and funding program managers to be engaged. V&V is extremely important but relies fundamentally on coupling with UQ.

4. Open source tools development and dissemination

Open source software tools refer to collaboratively developed testbed platforms that modelers and experimentalists can leverage to learn and apply standard V&V approaches. Such tools provide open access to "standard" or broadly accepted V&V/UQ practices; if properly applied, open source tools can deliver significant return on investment. One specific case is the development and dissemination of open source tools for UQ. This could be accomplished, for instance, by development and implementation of such software tools in existing design/lab-based courses in undergraduate and/or graduate curricula.

5. Challenge problems

Challenge problems, as outlined in section IV, also serve as a logical source of multidisciplinary training, through the participation of interdisciplinary teams involved in various modeling and measurement efforts. These challenge problems are typically conducted as a series of workshops/conferences where community members are invited to participate in teams. Thus, these challenge problems can serve as an excellent opportunity for practitioners and/or students to interact and learn from each other, as well as to learn good V&V/UQ practices from development of their own contributions to the challenge. Participants in such challenge problems are typically members of the existing computational modeling community, including graduate students and postdocs (as opposed to undergraduate students). It should be emphasized, however, that the challenge problems alone cannot address the fundamental need for formal education in these methods, as addressed in other sub-sections of this report.

6. Tutorials and texts

Web-based tutorials and published texts, both teaching texts and monographs, also are important resources for training and curricula support for understanding of V&V practices within the MOMS community. Beyond development of webinars, online videos offer another venue that educators are increasingly using, e.g., YouTube, to create and share content that provides discourse on key V&V skill sets. Such webinars, video tutorials, focused reports, and/or textbooks would be most applicable to existing computational modeling community members and graduate students, although there is also a place for such offerings in the early stages of exposure to these concepts for all early career stakeholders.

7. Mentoring

A key mechanism to transfer expert-level knowledge to students, as well as practicing scientists and engineers, is mentoring. Various mechanisms within organizations and/or fostered by professional societies could provide connections between established experts and V&V novices, allowing for the transfer of capability and knowledge. Such mentoring can be invaluable in making progress in incipient V&V exercises and working though the challenges that confront novice practitioners. An experienced mentor can act to augment progress and education in conjunction with any of the previous means and provide assistance in the practical application of V&V theory to real-world problems.

VIII. Final Comments and Call to Action

Predictive computational models associated with the mechanics of materials and structures (MOMS) can result in significant reductions in the cost and time to develop new materials, products, structures, and manufacturing procedures. But there is great concern whether such models can produce trustworthy results, with quantified levels of accuracy and uncertainty, for product/platform development and risk management during manufacturing and product lifetimes. Rigorous model verification and validation (V&V) and uncertainty quantification (UQ) are critical; yet, such robust practices are currently dramatically insufficient within MOMS-related communities. Thus, there is a strong need to accelerate the widespread adoption and implementation of V&V activities in MOMS computational modeling, to the point where these activities should become routine.

The major purpose of this study and report are to provide the motivation, framework, detailed recommendations, and some additional references/resources to help accelerate the widespread implementation of robust V&V/UQ techniques and practices within the mechanics of materials and structures and related communities. To address these goals, this report on *Accelerating the Broad Implementation of Verification and Validation in Computational Models of the Mechanics of Materials and Structures* captures and consolidates the ideas and outputs of internationally recognized technical experts from a variety of disciplines (see the Acknowledgments section). This study and report also built upon some key recommendations first identified in an earlier TMS workshop and report on this topic.[3]

After laying out the strong value proposition for V&V/UQ (section II), a framework and set of recommendations and detailed tasks for recommended practices for V&V/UQ of computational models associated with the mechanics of materials and structures are provided in section III. This is the most robust section of the report and might be considered a “field manual” of sorts for readers with little or no experience with robust V&V/UQ practices in MOMS-modeling, as well as for those who already have some experience but wish to dig deeper. To further help accelerate V&V and UQ in the MOMS realm, recommendations for challenge problems, sustainable funding programs, symposia/conferences, workshops, and/or future studies are provided in sections IV and V. Some recommended strategies for connecting modelers and experimentalists and improving training and curricula development are provided in sections VI and VII, respectively. All of these sections are meant to work together to help stakeholders and practitioners accelerate the widespread implementation of robust V&V/UQ techniques and practices within MOMS-based predictive computational modeling efforts.

Due to the nature of these practices, the activities recommended in this report should involve coordination among a number of disciplines and professional tracks – including, for example, both modelers and experimentalists in mechanical engineering, materials science and engineering, and civil engineering; solid mechanics experts; software/code developers; numerical and/or computational analysts; regulators and other customers of such predictive modeling efforts; program officers at funding organizations; and other resource gatekeepers in government, industry, or academia. Beyond just finding the content informative, readers of this report are challenged to use the information provided here to stimulate direct action. You could begin to act upon a number of the activities/actions recommended here almost immediately. You also are encouraged to use this information to impart knowledge to others and to stimulate the development of additional ideas and activities that will contribute to the adoption of robust V&V practices in computational models of the mechanics of materials and structures. Some general next steps could include: (1) identifying specific V&V and UQ activities that you and your colleagues could address, which would be most relevant to your personal and/or organizational goals and activities; (2) sketching out a detailed action plan and timeline for the activities discussed herein; and (3) taking concrete steps to initiate these activities.

Beyond just finding the content informative, readers of this report are challenged to use the information provided here to stimulate direct action. You could begin to act upon a number of the activities/actions recommended here almost immediately.

Our desire is that the readers of this report will act both promptly and in a sustained, long-term fashion on the activities and/or practices recommended here. The specific recommendations and activities should not be viewed as all-inclusive and could be used to initiate conversations that determine what would be appropriate for you and your organization. On the other hand, engaging in all or even a majority of the detailed tasks and recommendations provided in section III should not be viewed as a necessity for all practitioners reading this report, as it is recognized that readers will be at various levels of experience, expertise, and engagement in V&V/UQ practices. This report was thus written with the intent that each individual across this wide spectrum - from novices with almost no V&V/UQ knowledge to those that have a great deal of experience - can take away the information and practices that are of greatest benefit to them, personally and/or organizationally.

The potential is thus great for a wide variety of stakeholders who read this report to make rapid progress (as well as foundational, longer-term contributions) toward implementing more robust V&V/UQ practices in computational modeling efforts. Such activity is vitally needed in order to bring to fruition the great potential of predictive models in supporting the development of advanced new materials, components, and structures.

References

1. L. E. Schwer, “An overview of the PTC 60/V&V 10: guide for verification and validation in computational solid mechanics,” *Engineering with Computers*, vol. 23, no. 4, pp. 245–252, Dec. 2007, doi: 10.1007/s00366-007-0072-z.

2. National Research Council, *Assessing the Reliability of Complex Models: Mathematical and Statistical Foundations of Verification, Validation, and Uncertainty Quantification*. National Academies Press, 2012.

3. The Minerals, Metals, and Materials Society (TMS), *Verification & Validation of Computational Models Associated with the Mechanics of Materials*. Warrendale, PA: TMS, 2019.

4. I. Babuska and J. T. Oden, “Verification and validation in computational engineering and science: basic concepts,” *Computer Methods in Applied Mechanics and Engineering*, vol. 193, no. 36–38, pp. 4057–4066, Sep. 2004, doi: 10.1016/j.cma.2004.03.002.

5. White House Office of Science and Technology Policy, “Materials Genome Initiative for Global Competitiveness,” Jun. 2011. [Online]. Available: https://www.whitehouse.gov/sites/default/files/microsites/ostp/materials_genome_initiative-final.pdf.

6. National Science and Technology Council - Committee on Technology Subcommittee on the Materials Genome Initiative, “Materials Genome Initiative Strategic Plan,” Dec. 2014. [Online]. Available: https://mgi.nist.gov/sites/default/files/factsheet/mgi_strategic_plan_-_dec_2014.pdf.

7. R. G. Sargent, "Verification and validation of simulation models," *Journal of Simulation*, vol. 7, no. 1, pp. 12–24, Feb. 2013, doi: 10.1057/jos.2012.20.

8. The American Society of Mechanical Engineers, *Standard for Verification and Validation in Computational Solid Mechanics*. New York, NY: ASME, 2020.

9. C. J. Freitas, "Standards and Methods for Verification, Validation, and Uncertainty Assessments in Modeling and Simulation," *J. Verif. Valid. Uncert*, vol. 5, no. 2, Jun. 2020, doi:10.1115/1.4047274.

10. C. J. Roy and W. L. Oberkampf, "A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing," *Computer Methods in Applied Mechanics and Engineering*, vol. 200, no. 25, pp. 2131–2144, Jun. 2011, doi: 10.1016/j.cma.2011.03.016.

11. B. L. Boyce *et al.*, "The Sandia Fracture Challenge: blind round robin predictions of ductile tearing," *Int J Fract*, vol. 186, no. 1, pp. 5–68, Mar. 2014, doi: 10.1007/s10704-013-9904-6.

12. B. L. Boyce *et al.*, "The Second Sandia Fracture Challenge : Predictions of Ductile Failure Under Quasi-static and Moderate-rate Dynamic Loading," *International Journal of Fracture*, vol. 198, pp. 5–100, 2016, doi: 10.1007/s10704-016-0089-7.

13. S. L. B. Kramer *et al.*, "The third Sandia fracture challenge: predictions of ductile fracture in additively manufactured metal," *Int J Fract*, vol. 218, no. 1, pp. 5–61, Jul. 2019, doi: 10.1007/s10704-019-00361-1.

14. K. Pack, M. Luo, and T. Wierzbicki, "Sandia Fracture Challenge: blind prediction and full calibration to enhance fracture predictability," *Int J Fract*, vol. 186, no. 1, pp. 155–175, Mar. 2014, doi: 10.1007/s10704-013-9923-3.

15. W. L. Oberkampf, T. G. Trucano, and C. Hirsch, "Verification, validation, and predictive capability in computational engineering and physics," *Appl. Mech. Rev*, vol. 57, no. 5, pp. 345–384, Dec. 2004, doi: 10.1115/1.1767847.

16. W. L. Oberkampf and T. G. Trucano, "Verification and validation in computational fluid dynamics," *Progress in Aerospace Sciences*, vol. 38, no. 3, pp. 209–272, Apr. 2002, doi: 10.1016/S0376-0421(02)00005-2.

17. J. J. Caro, A. H. Briggs, U. Siebert, and K. M. Kuntz, "Modeling Good Research Practices—Overview: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force–1," *Medical Decision Making*, Sep. 2012, doi: 10.1177/0272989X12454577.

18. The Minerals Metals & Materials Society (TMS), *Modeling Across Scales: A Roadmapping Study for Connecting Materials Models and Simulations Across Length and Time Scales*. Warrendale, PA: TMS, 2015.

19. B. Cowles, D. Backman, and R. Dutton, "Verification and Validation of ICME Methods and Models for Aerospace Applications," *Integrating Materials*, vol. 1, no. 1, p. 2, Dec. 2012, doi: 10.1186/2193-9772-1-2.

20. J. Allison, M. Li, C. Wolverton, and X. Su, "Virtual aluminum castings: An industrial application of ICME," *JOM*, vol. 58, no. 11, pp. 28–35, Nov. 2006, doi: 10.1007/s11837-006-0224-4.

21. The Minerals Metals & Materials Society (TMS), *Integrated Computational Materials Engineering (ICME): Integrating ICME in The Automotive, Aerospace, and Maritime Industries*. Warrendale, PA: TMS, 2013.

22. National Science Board, "Publications Outputs: U.S. Trends and International Comparisons." Accessed: Jun. 09, 2020. [Online]. Available: https://ncses.nsf.gov/pubs/nsb20206/executive-summary.

23. J. A. Ang, T. G. Trucano, and D. R. Luginbuhl, "Confidence in ASCI scientific simulations," *Hawaii international conference on system sciences (HICSS 31), Kohala Coast, HI (United States), 6-9 Jan 1998*, Jun. 01, 1998. https://digital.library.unt.edu/ark:/67531/metadc704083/ (accessed Sep. 16, 2020).

24. The American Society of Mechanical Engineers, *Guide for Verification and Validation in Computational Solid Mechanics*. New York, NY: ASME, 2006.

25. T. G. Trucano, L. P. Swiler, T. Igusa, W. L. Oberkampf, and M. Pilch, "Calibration, validation, and sensitivity analysis: What's what," *Reliability Engineering & System Safety*, vol. 91, no. 10, pp. 1331–1357, Oct. 2006, doi: 10.1016/j.ress.2005.11.031.

26. Y. Wang and D. L. McDowell, "Uncertainty quantification in materials modeling," in *Uncertainty Quantification in Multiscale Materials Modeling*, Y. Wang and D. L. McDowell, Eds. Woodhead Publishing, 2020, pp. 1–40.

27. B. Szabo and R. Actis, "Simulation governance: Technical requirements for mechanical design," in *Computer methods in applied mechanics and engineering*, 2012, vol. 249–52, pp. 158–168, Accessed: Sep. 16, 2020. [Online]. Available: http://pascal-francis.inist.fr/vibad/index.php?action=getRecordDetail&idt=26842870.

28. B. M. Wilson and A. Koskelo, "A Practical Validation Assessment Workflow," *J. Verif. Valid. Uncert*, vol. 5, no. 1, Mar. 2020, doi: 10.1115/1.4046787.

29. P. J. Roache, “Code Verification by the Method of Manufactured Solutions,” *J. Fluids Eng*, vol. 124, no. 1, pp. 4–10, Mar. 2002, doi: 10.1115/1.1436090.

30. P. J. Roache, “The Method of Manufactured Solutions for Code Verification,” in *Computer Simulation Validation: Fundamental Concepts, Methodological Frameworks, and Philosophical Perspectives*, C. Beisbart and N. J. Saam, Eds. Cham: Springer International Publishing, 2019, pp. 295–318.

31. K. Salari and P. Knupp, “Code Verification by the Method of Manufactured Solutions,” 2000, [Online]. Available: https://prod-ng.sandia.gov/techlib-noauth/access-control.cgi/2000/001444.pdf.

32. D. W. Trott and M. K. Gobbert, “Conducting Finite Element Convergence Studies using COMSOL 4.0,” *Proceedings of the COMSOL Conference*, p. 6, 2010.

33. P. J. Roache, *Verification and validation in computational science and engineering*. Albuquerque, NM: Hermosa, 1998.

34. C. J. Roy, “Review of code and solution verification procedures for computational simulation,” *Journal of Computational Physics*, vol. 205, no. 1, pp. 131–156, May 2005, doi: 10.1016/j.jcp.2004.10.036.

35. L. Eça, G. Vaz, and M. Hoekstra, “A verification and validation exercise for the flow over a backward facing step,” presented at the European Conference on Computational Fluid Dynamics, ECCOMAS CFD 2010, Jun. 2010.

36. W. Rider, W. Witkowski, J. R. Kamm, and T. Wildey, “Robust verification analysis,” *Journal of Computational Physics*, no. 307, pp. 146–163, 2016.

37. J. C. for G. in Metrology, “Evaluation of measurement data—Guide to the expression of uncertainty in measurement,” *JCGM*, vol. 100, no. 2008, pp. 1–116, 2008.

38. A. G. González, M. A. Herrador, and A. G. Asuero, *A Practical Way to ISO/GUM Measurement Uncertainty for Analytical Assays Including In-House Validation Data*. IntechOpen, 2018.

39. H. W. Coleman and W. G. Steele, *Experimentation and Uncertainty Analysis for Engineers*, 2nd Edition. New York, NY: Wiley-Interscience, 1999.

40. A. Al-Refaie and N. Bata, “Evaluating measurement and process capabilities by GR&R with four quality measures,” *Measurement*, vol. 43, no. 6, pp. 842–851, Jul. 2010, doi: 10.1016/j.measurement.2010.02.016.

41. J. Pan, “Determination of the optimal allocation of parameters for gauge repeatability and reproducibility study,” *International Journal of Quality & Reliability Management*, vol. 21, no. 6, pp. 672–682, Jan. 2004, doi: 10.1108/02656710410542061.

42. W. D. Kappele and J. Raffaldi, "Gage R&R for Destructive Measurement Systems," *Quality Magazine*, vol. 5, pp. 32–34, 2010.

43. M. Pilch, T. G. Trucano, and J. C. Helton, "Ideas underlying the Quantification of Margins and Uncertainties," *Reliability Engineering & System Safety*, vol. 96, no. 9, pp. 965–975, Sep. 2011, doi: 10.1016/j.ress.2011.03.016.

44. W. L. Oberkampf and C. J. Roy, *Verification and Validation in Scientific Computing*, 1st Edition. New York: Cambridge University Press, 2010.

45. T. J. R. Hughes, G. R. Feijóo, L. Mazzei, and J.-B. Quincy, "The variational multiscale method—a paradigm for computational mechanics," *Computer Methods in Applied Mechanics and Engineering*, vol. 166, no. 1, pp. 3–24, Nov. 1998, doi: 10.1016/S0045-7825(98)00079-6.

46. A. Masud, T. J. Truster, and L. A. Bergman, "A variational multiscale a posteriori error estimation method for mixed form of nearly incompressible elasticity," *Computer Methods in Applied Mechanics and Engineering*, vol. 200, no. 47, pp. 3453–3481, Nov. 2011, doi: 10.1016/j.cma.2011.07.012.

47. A. Masud and T. J. Truster, "A framework for residual-based stabilization of incompressible finite elasticity: Stabilized formulations and $\bar{F}$ methods for linear triangles and tetrahedra," *Computer Methods in Applied Mechanics and Engineering*, vol. 267, pp. 359–399, Dec. 2013, doi: 10.1016/j.cma.2013.08.010.

48. The Minerals Metals & Materials Society (TMS), "Additive Manufacturing Benchmarks 2018 (AM-Bench 2018)." www.tms.org/AMBench2018 (accessed September 09, 2020).

49. M. W. Barclift and C. B. Williams, "Examining Variability in the Mechanical Properties of Parts Manufactured Via PolyJet Direct 3D Printing," in *International. Solid Freeform Fabrication Symposium Proceedings*, Austin, TX, Jan. 2012, p. 15.

50. J. J. Lewandowski and M. Seifi, "Metal Additive Manufacturing: A Review of Mechanical Properties," *Annual Review of Materials Research*, vol. 46, no. 1, pp. 151–186, 2016, doi: 10.1146/annurev-matsci-070115-032024.

51. A. du Plessis, I. Yadroitsava, and I. Yadroitsev, "Effects of defects on mechanical properties in metal additive manufacturing: A review focusing on X-ray tomography insights," *Materials & Design*, vol. 187, p. 108385, Feb. 2020, doi: 10.1016/j.matdes.2019.108385.

52. T. Gatsos, K. A. Elsayed, Y. Zhai, and D. A. Lados, "Review on Computational Modeling of Process–Microstructure–Property Relationships in Metal Additive Manufacturing," *JOM*, vol. 72, no. 1, pp. 403–419, Jan. 2020, doi: 10.1007/s11837-019-03913-x.

53. The Minerals Metals & Materials Society (TMS), *Harnessing Materials Innovations to Support Next Generation Manufacturing Technologies*. Warrendale, PA: TMS, 2018.

54. H. W. Lee and S. W. Kang, "The Relationship between Residual Stresses and Transverse Weld Cracks in Thick Steel Plate," *Welding Journal*, vol. 82, pp. 225–230, 2003.

55. G. Spanos, D. W. Moon, R. W. Fonda, E. S. K. Menon, and A. G. Fox, "Microstructural, compositional, and microhardness variations across a gas-metal arc weldment made with an ultralow-carbon consumable," *Metall and Mat Trans A*, vol. 32, no. 12, pp. 3043–3054, Dec. 2001, doi: 10.1007/s11661-001-0179-6.

56. D. W. Moon, R. W. Fonda, and G. Spanos, "Microstructure and Property Evolution in HSLA-100 Steel Welds Fabricated with New Ultra Low Carbon Filler Metals," *The Weldong Journal, Research Supplement*, vol. 79, no. 10, pp. 278–285, Oct. 2000.

57. H. Guo, J. Hu, and H. L. Tsai, "Three-Dimensional Modeling of Gas Metal Arc Welding of Aluminum Alloys," *J. Manuf. Sci. Eng*, vol. 132, no. 2, Apr. 2010, doi: 10.1115/1.4001479.

58. E. A. Metzbower, G. Spanos, R. W. Fonda, and R. A. Vandermeer, "Prediction of microstructure and properties in the heat affected zone of HSLA-100 weldments," *Science and Technology of Welding and Joining*, vol. 2, no. 1, pp. 27–32, Feb. 1997, doi: 10.1179/stw.1997.2.1.27.

59. E. A. Metzbower, G. Spanos, R. W. Fonda, and R. A. Vandermeer, "Microstructures in the Heat-Affected Zones Based on Dilatometer Data," *NIST special publication*, no. 923, pp. 298–309, 1997.

60. F. Roters, P. Eisenlohr, L. Hantcherli, D. D. Tjahjanto, T. R. Bieler, and D. Raabe, "Overview of constitutive laws, kinematics, homogenization and multiscale methods in crystal plasticity finite-element modeling: Theory, experiments, applications," *Acta Materialia*, vol. 58, no. 4, pp. 1152–1211, Feb. 2010, doi: 10.1016/j.actamat.2009.10.058.

61. S. Papanikolaou, Y. Cui, and N. Ghoniem, "Avalanches and plastic flow in crystal plasticity: an overview," *Modelling Simul. Mater. Sci. Eng.*, vol. 26, no. 1, p. 013001, Dec. 2017, doi: 10.1088/1361-651X/aa97ad.

62. B. Pan, K. Qian, H. Xie, and A. Asundi, "Two-dimensional digital image correlation for in-plane displacement and strain measurement: a review," *Meas. Sci. Technol.*, vol. 20, no. 6, p. 062001, Apr. 2009, doi: 10.1088/0957-0233/20/6/062001.

63. T. A. Laursen, *Computational Contact and Impact Mechanics: Fundamentals of Modeling Interfacial Phenomena in Nonlinear Finite Element Analysis*. Springer Science & Business Media, 2013.

64. G. Guo, Q. Wang, and Y. Rong, "A Brief Review of Precipitate Hardening Models for Aluminum Alloys," presented at The 2nd World Congress on Integrated Computational Materials Engineering (ICME), 2013.

65. M. R. Ahmadi, E. Povoden-Karadeniz, K. I. Öksüz, A. Falahati, and E. Kozeschnik, "A model for precipitation strengthening in multi-particle systems," *Computational Materials Science*, vol. 91, pp. 173–186, Aug. 2014, doi: 10.1016/j.commatsci.2014.04.025.

66. G. H. Teichert, A. R. Natarajan, A. Van der Ven, and K. Garikipati, "Machine learning materials physics: Integrable deep neural networks enable scale bridging by learning free energy functions," *Computer Methods in Applied Mechanics and Engineering*, vol. 353, pp. 201–216, Aug. 2019, doi: 10.1016/j.cma.2019.05.019.

67. G. H. Teichert, A. R. Natarajan, A. Van der Ven, and K. Garikipati, "Scale bridging materials physics: Active learning workflows and integrable deep neural networks for free energy function representations in alloys," *Computer Methods in Applied Mechanics and Engineering*, vol. 371, p. 113281, Nov. 2020, doi: 10.1016/j.cma.2020.113281.

68. J. Mattis, "Summary of the 2018 National Defense Strategy," p. 14.

69. A. W. Sazali, C. R. Raduan, U. Jegak, and A. Haslinda, "Evolution and development of technology transfer models and the influence of knowledge-based view and organizational learning on technology transfer," *Research Journal of International Studies*, vol. 12, pp. 79–91, 2009.

70. A. Singh and G. Aggarwal, "Technology Transfer Introduction, Facts and Models," *International Journal of Pharma World Research*, vol. 1, no. 2, pp. 1–8, 2010.

71. The Minerals Metals & Materials Society (TMS), *Building a Materials Data Infrastructure: Opening New Pathways to Discovery and Innovation in Science and Engineering*. Pittsburgh, PA: TMS, 2017.

72. The Minerals Metals & Materials Society (TMS), *Creating the Next-Generation Materials Genome Initiative Workforce*. Pittsburgh, PA: TMS, 2019.

73. "The Minerals Metals & Materials Society (TMS), *Metamorphic Manufacturing: Shaping the Future of On-Demand Components*. Pittsburgh, PA: TMS, 2019.

Appendix: Glossary

Terms	Meanings
acceptable agreement with experiment	the simulation results agree with experimental measurements, including uncertainty, within the stated accuracy requirements; provides the primary basis for the assertion that the model can accurately predict future scenarios; demonstrable "acceptable agreement" does not imply a lack of uncertainty - typical questions asked around the topic of "acceptable agreement" include: • What is the evidence package used for decision-making? • How far is the model-based extrapolation beyond the validation database? • What is the estimated total uncertainty in the predictions? • What are the consequences if the system fails to perform as predicted?
accuracy requirements	set of requirements or metrics for the accuracy of a simulation or experimental measurement, for specified conditions of interest

adequacy requirements	criteria set by the user of the simulation results or the regulator on whether the model sufficiently represents reality; sufficiency of model fidelity for all of the intended use conditions of interest
application domain	indicates the range of physical conditions where predictions are needed from the model for the application of interest; the domain that comprises the physical conditions for the real/ physical system of interest; the application domain is commonly specified in terms of the range of values of all input parameters of the model; uncertainties in the application domain are not necessarily related to uncertainties that exist in laboratory experiments
calibration	refers to the tuning or updating of model input parameters using an experimental dataset; in calibration the model is assumed to be correct and model parameters are defined to achieve agreement between predicted and experimentally observed values; calibration of parameters requires the solution of an inverse model problem to estimate the parameters
challenge problem	typically conducted as a series of workshops where the V&V community members are invited to participate in teams to demonstrate and stress-test V&V approaches; V&V/UQ challenge problems typically focus especially on one aspect, such as code verification, solution verification, model validation, or uncertainty quantification in the MOMS community
confidence (statistical)	difference between an estimated statistic of a set of samples, e.g., the mean, and the true value of the statistic
error	deviation of a predicted quantity from its true value
evidence package	evidence of verification and validation procedures; for instance, the evidence package for the verification process includes input and output from the code, analysis of the solutions, and subsequent convergence analysis
gage repeatability and reproducibility (GR&R)	process of reducing measurement variability and improving the repeatability and reproducibility of experimental measurements

hierarchical validation	seeks to disassemble the full system into tiers or building blocks with lower levels of physical complexity; identifies a range of required experiments on specific subsystems or subcomponents; an objective in hierarchical validation is estimation of various sources of uncertainties and improvement of models at lower levels in the hierarchy and subsequently quantification of those uncertainties at higher levels, i.e., with increasing complexity
materials qualification	the practice of generating a statistical basis for material acceptance and quality control to verify the materials performance for a given application
method of manufactured solutions (MMS)	a tool for developing exact solutions for simulations without traditional analytical solutions; a solution is determined and then a source term is derived that forces the system to have that solution
model	conceptual/mathematical/numerical description of a specific physical scenario, including geometrical, material, initial, and boundary data
model - conceptual model	collection of assumptions, algorithms, relationships, approximations, and data that describe the reality of interest from which the mathematical model can be constructed
model - computational model	the discretized realization of the mathematical model in the computer software
model - mathematical model	mathematical equations that describe the system of interest, including the initial conditions, geometric features, physical modeling parameters of the system boundary conditions, and system excitation conditions needed to describe the conceptual model of the system of interest
predictive capability	refers to the credibility of how much the model is capable of predicting; predictive capability includes all relevant sources of uncertainty impacting the simulation result

quantities of interest (QoI)	Quantities of Interest (QoI) are physical entities or features related to the target outputs and prediction goals of the mathematical model, the values of which are of interest to stakeholders, typically because they inform decisions; QoI can be experimentally measured quantities, model input quantities, or system response quantities of interest; some examples of QoI could include: (1) thermal conductivity, (2) maximum Von Mises stress, and/or (3) the ratio between the maximum temperature in a part and the material's melting temperature.
uncertainty	imperfectly known information concerning an object or issue; potential deficiency in any phase or activity of the modeling or experimentation process that is due to inherent variability (irreducible uncertainty) or lack of knowledge (reducible uncertainty)
uncertainty - aleatoric uncertainty	uncertainty that is due to inherent randomness or variability in any quantity of interest
uncertainty - epistemic uncertainty	uncertainty that is due to our lack of knowledge of any quantity of interest, system of interest, of scenario of interest
uncertainty - model-form uncertainty	uncertainty that is due to assumptions and approximations made in the formulation of the mathematical model
uncertainty - parametric uncertainty	uncertainty that is due to the variations of the input variables or parameters
validation domain	indicates the domain within the input parameter space where model validation experiments have been conducted; the convex hull of the input parameter space of model validation experiments
validation experiments	experiments performed to generate experimental data required for the assessment of model accuracy; these experimental data include not only system response quantities of interest, but also all model input data needed to simulate the experiment

verification - code verification	determination of whether the numerical solution from the code accurately represents the exact solution of the mathematical model; comparison of the exact solution to the mathematical model to software output; test of discrete and iterative convergence of all algorithms and software steps that produce the numerical solution; test of the code solution against a known analytic (or trusted) solution; collection of evidence to establish confidence in the ability of the mathematical models and solution algorithms to function properly
verification - solution verification/ calculation verification	estimation of the numerical error resident in a solution based on the discretization of the mathematical model; numerical estimation of calculation error as a function of discretization and iterative features in the computer software; collection of evidence to establish confidence in the accuracy of the discretized solution of the mathematical model

The Minerals, Metals & Materials Society

Promoting the global science and engineering professions concerned with minerals, metals and materials