VOLUME 8 | NUMBER 1 | 2024

DSIAC JOURNAL

PAGE 22

A SURROGATE MODEL TO QUANTIFY UNCERTAINTY

IN THERMAL PROTECTION SYSTEMS FOR HYPERSONIC WEAPONS

DSIAC JOURNAL

Volume 8 // Number 1 // 2024

The DSIAC Journal is a publication of the Defense Systems Information Analysis Center (DSIAC). DSIAC is a DoD Information Analysis Center (IAC) sponsored by the Defense Technical Information Center (DTIC) with policy oversight provided by the Office of the Under Secretary of Defense (OUSD) for Research and Engineering (R&E). DSIAC is operated by the SURVICE Engineering Company.

ISSN 2471-3392 (Print) // ISSN 2471-3406 (Online)

Distribution Statement A:
Approved for public release; distribution is unlimited.

On the Cover:
Digital Art Rendering of Falcon HTV-2 *(Source: Darpa)*.

Defense Systems Information Analysis Center

ABOUT DSIAC

Who We Are

A DoD Information Analysis Center comprised of scientists, engineers, researchers, analysts, and information specialists.

What We Do

Generate, collect, research, analyze, synthesize, and disseminate scientific and technical information (STI) to DoD and federal government users and industry contractors.

Why Our Services

To eliminate redundancy, foster collaboration, and stimulate innovation.

DSIAC SERVICES

Subject Matter Expert (SME) Connections

Access to a network of experts with expertise across our technical focus areas.

Technical Inquiries (TIs)

Up to 4 hours of FREE research using vast DoD information resources and our extensive network of SMEs.

Specialized Task Orders

Research and analysis services to solve our customer's toughest scientific and technical problems.

Webinars & Events

Our webinars feature a technical presentation from a SME in one of our focus areas. We also offer key technical conferences and forums for the science and technology community.

STI Collection

Our knowledge management team collects and uploads all pertinent STI into DTIC's Research & Engineering Gateway.

Information Research Products

The Defense Systems Digest, state-of-the-art reports, journals, TI response reports, and more available on our website.

CONTACT DSIAC

IAC Program Management Office

8725 John J. Kingman Road
Fort Belvoir, VA 22060
Office: 571.448.9753

DSIAC Headquarters

4695 Millennium Drive
Belcamp, MD 21017-1505
Office: 443.360.4600
Fax: 410.272.6763
Email: contact@dsiac.org

DSIAC Technical Project Lead

Brian Benesch
4695 Millennium Drive
Belcamp, MD 21017-1505
Office: 443.360.4600

FEATURED ARTICLE

A SURROGATE MODEL TO QUANTIFY UNCERTAINTY IN THERMAL PROTECTION SYSTEMS FOR HYPERSONIC WEAPONS

By Eric A. Walker, Jason Sun, and James Chen

Modeling and simulation are key for the iterative development of thermal protection systems for hypersonic weapons. In this work, the temperature-dependent flexural strength of α-SiC ceramic is predicted given Young's modulus, Poisson's ratio, and temperature. An artificial neural network surrogate model is created to retain property-performance prediction while increasing computation speed.

IN THIS ISSUE

WHAT MAKES A SIMULATION CREDIBLE?

Cost-Effective VV&A in the Systems Engineering Process

BY DAVID H. HALL AND DAVID J. TURNER (PHOTO SOURCE: 123RF.COM AND CANVA)

INTRODUCTION

The U.S. Naval Air Warfare Center Aircraft Division (NAWCAD) Verification, Validation, and Accreditation (VV&A) Branch has developed and is executing a cost-effective, risk-based VV&A process for models and simulations (M&S) used to support the U.S. Department of Defense (DoD). The original version of the process was created more than 25 years ago by the predecessor to the VV&A Branch, the Joint Accreditation Support Activity, and has been further developed, used, and refined over the last 20 years in support of a wide variety of M&S domains and military systems. It is a systematic and straightforward way to determine whether a proposed M&S has the credibility to support its intended uses. This risk-based VV&A process equally applies to cases where accreditation is required and cases where formal accreditation is not required but verification and validation (V&V) is needed. It also applies to any M&S and test facilities that include live, virtual, and constructive simulations.

WHAT IS CREDIBILITY?

Based on over 30 years of working in the field of M&S credibility and discussing what it means to various scientists, engineers, and mathematicians using the results of M&S for decision-making, the authors have concluded that M&S credibility is a function of the following three factors:

1. Capability – the functions it models and the level of detail with which it is modeled should support anticipated uses.
2. Accuracy – how accurate it must be should depend on the risks involved if the answers are incorrect.
3. Usability – the extent of available user support should ensure it is not misused.

Any robust assessment of M&S credibility must consider not only accuracy but capability and usability. Capability is the characteristic that ties the M&S to the problem; it describes what the M&S needs to do to support the intended use. Accuracy describes how well the M&S solves the problem in terms of three elements—software accuracy, input/embedded data accuracy, and output accuracy. Usability ties the M&S to a useful solution by ensuring that it will not be misused. Credibility should be defined in terms of those three characteristics as follows:

> *Any robust assessment of M&S credibility must consider not only accuracy but capability and usability.*

M&S Credibility: The M&S has sufficient capability, accuracy (software, data, and output), and usability to support the intended use.

NAVAL AIR SYSTEMS COMMAND (NAVAIR) VV&A PROCESS

The risk-based VV&A process starts with defining the intended uses of the M&S to support program decision-making. It continues with a detailed analysis of what is required for the M&S to satisfy those intended uses, how to demonstrate that those requirements have been met (or not), and what metrics will be used to measure the M&S against those requirements. A final check on the suitability of the M&S for the intended use is a risk assessment: What are the risks of using the M&S for the intended purpose given all that is known about its credibility once the VV&A effort is completed?

Figure 1 depicts an overview of steps in the process, with responsibilities assigned to the various organizations involved. The first and possibly hardest step in the VV&A process is to thoroughly articulate the intended use of the M&S for the application at hand (what questions will be answered using M&S outputs and how). The next step is to develop M&S and credibility requirements and acceptance criteria based on

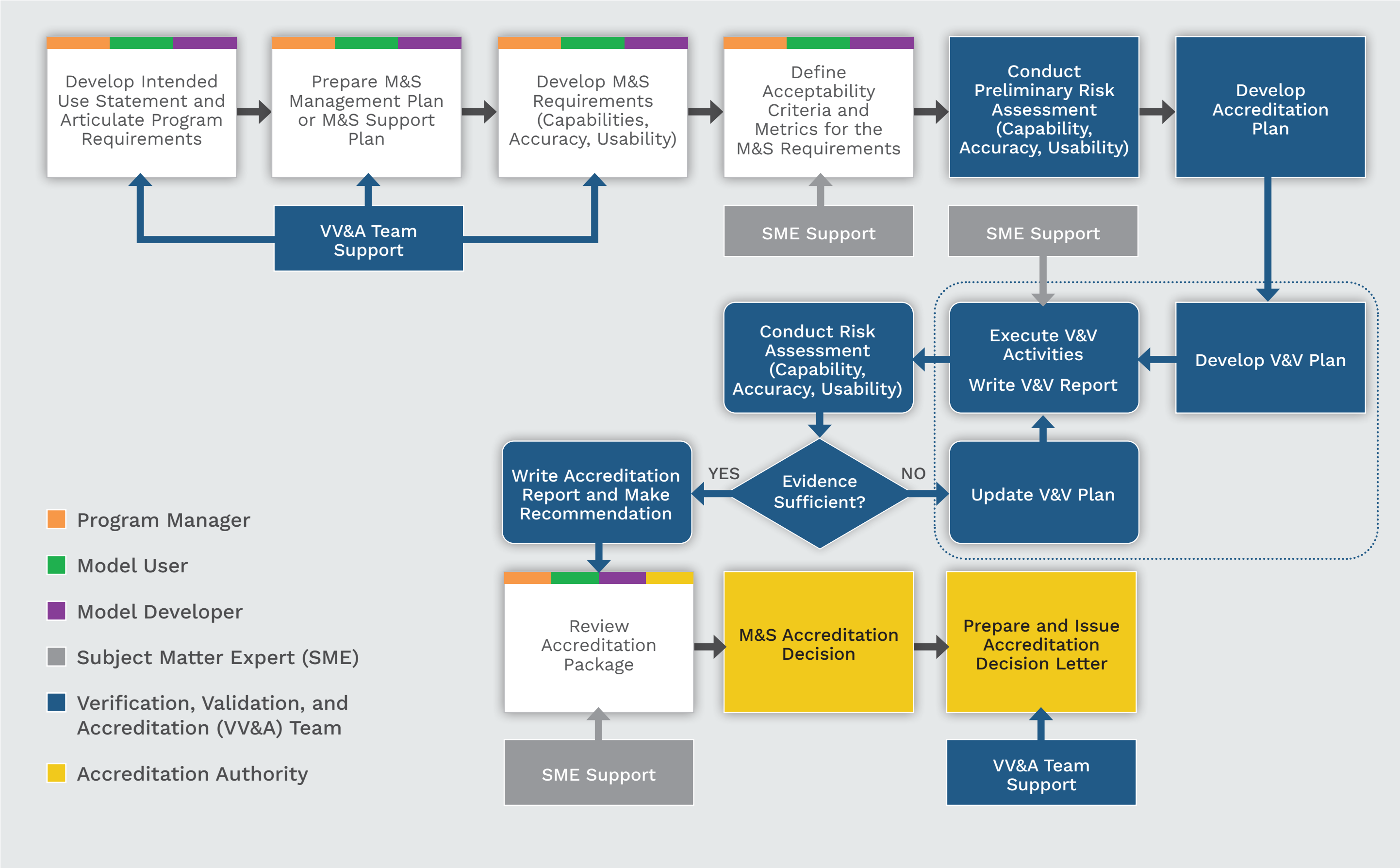


Figure 1. NAVAIR Risk-Based VV&A Process Overview *(Source: NAVAIR VV&A Branch).*

that specific intended use statement (SIUS) (what will the M&S need to do, how accurate will it be, and what information will the accreditor need to see to decide whether to accept it or not). The next step is to assess the risk of using M&S results for the intended use; the risk assessment identifies gaps in credibility that need filled and builds an accreditation case for the M&S. Those gaps form the basis of V&V and accreditation plans. Subject matter expert (SME) reviews of the information elicit SME accreditation recommendations and face validation results. The accreditation authority reviews the accreditation case and any residual risks before deciding to accept the risk of using the M&S, rejecting it, or accepting it with restrictions and/or workarounds. In the next few sections, an example for a six degree of freedom (6-DOF) flight simulation of a long-endurance, unmanned aerial vehicle will be presented.

SIUS

The purpose of the SIUS is to state the program's goals for the M&S concisely and completely, describe a potential M&S user's needs and questions, and explain how M&S might help meet those needs. An SIUS must be developed in enough detail so that accreditation requirements can be determined (general statements of intended use are insufficient) [1]:

> Carefully define the *specific* issues to be investigated by the study and the measures of performance that will be used for evaluation. Models are not universally valid but are designed for specific purposes…A great model for the wrong problem will never be used…

Table 1 shows an example from an SIUS for a 6-DOF flight simulation.

Table 1. Air Vehicle 6-DOF Flight Simulation SIUS

GENERAL INTENDED USE	QUESTIONS BEING ADDRESSED	SPECIFIC APPLICATION	M&S OUTPUTS USED
Verify specifications for flight dynamics and performance and failure mode requirements.	What are the flight characteristics of the vehicle in various conditions, and do they meet the specifications and requirements?	Present results from the aerodynamics model, actuator model, and equations of motion, mass properties, and engine model to support analysis in defining the airplane normal for each applicable flight phase and flight stability analysis.	Specific fuel consumption, endurance and time on station, engine-out responses and characteristics.
	What are the gain and phase margins? Does the Flight Control System (FCS) allow commands to drive the vehicle outside the flight envelope?	Develop gain and phase margins to ensure that the FCS compensates for errors, time delays, asymmetric flight from engine and fuel management, and loss of engine. Evaluate FCS oscillations, failures, control surface failures, longitudinal motions, and asymmetric center of gravity shifts.	Gain and phase margins, specific fuel consumption, endurance and time on station, range response to environmental effects, engine-out responses and characteristics.
Verify specifications for flight dynamics and performance related to atmospheric disturbance.	Does the vehicle meet performance specifications for response to atmospheric disturbance?	Examine interactions between the FCS and structural modes of the vehicle transitioning through atmospheric conditions.	Specific fuel consumption, endurance and time on station, range response to environmental effects, engine-out responses and characteristics.

M&S Accreditation Requirements, Acceptability Criteria, and Metrics

Whether an M&S is credible and hence acceptable for an application (intended use) is determined by how well it meets the requirements of that intended use. For VV&A, requirements, acceptability criteria, and metrics/measures are defined as follows:

- M&S Requirements: What features and characteristics does M&S need to support the intended use?
- Acceptability Criteria: What quantitative or qualitative properties must M&S have to meet the requirements for intended use?
- Metrics/Measures: How will it be determined whether the acceptability criteria are met?

The best way to define M&S requirements, acceptability criteria, and metrics is in terms of the three components that define M&S credibility—capability, accuracy (software, data, and outputs), and usability. How well the M&S meets those requirements must be determined by an assessment method, with criteria identified as to how the user will decide if it passes or fails. These three key credibility components determine the M&S features needed to satisfy the intended uses. Note that output accuracy is a comparison between M&S outputs and

> ***The best way to define M&S requirements, acceptability criteria, and metrics is in terms of the three components that define M&S credibility—capability, accuracy, and usability.***

a representation of the real world; the real world can be represented in three ways—benchmarking (comparing with another M&S of known credibility), face validation (comparing with SME opinions on how the system being simulated behaves in the real world), and results validation (comparing with test data).

Developing the details of the requirements, the acceptability criteria for each requirement, and the metrics and measures to evaluate the M&S against those acceptability criteria requires working with M&S developers and users (such as the M&S integrated product team [IPT] lead). Requirements for VV&A activities are developed from the SIUS, the acceptability criteria, and metrics—i.e., how much credibility information is required to demonstrate whether the M&S meets the intended use depends on what is required to show if it meets the requirements (acceptability criteria), how it is measured (metrics/measures), and the risks associated with incorrect results. Table 2 shows a summary of the types of criteria and metrics used. Table 3 shows an example of part of a matrix of M&S requirements (acceptability criteria and metrics for the same air vehicle [AV] 6-DOF flight simulation previously discussed).

Risk Assessment

Based on the available credibility information (including V&V results), a formal risk assessment process is used to evaluate the risk associated with making an accreditation decision. A set of 10 characteristics of M&S has been developed under the three categories of M&S credibility (capability, accuracy, and usability). The M&S risk is evaluated for the intended uses by reviewing the available credibility evidence in each of those 10 areas. Table 4 shows these characteristics and the criteria for a "green" rating in each. Table 5 illustrates the results of the risk assessment of the 6-DOF flight simulation as an example of applying the risk assessment process; the overall risk was assessed as "moderate." An operational risk assessment is also conducted to determine the effects of errors on various parameters important to the intended use.

Plan V&V to Reduce Risk

Preliminary risk assessments, which result in an initial "gap assessment" of the available credibility information on the M&S, are conducted to focus V&V-related efforts on the gaps identified and the risks to the intended uses associated with those gaps. This process leads to identifying requirements for additional information to be collected or generated to reduce those risks. These information requirements are then compared with any additional available information, and a list of credibility "shortfalls" is compiled. Each element of this list is then evaluated for its impact on risk. Unmet requirements for simulation credibility that have acceptable (i.e., low risk) workarounds are removed from the list. Unmet requirements for simulation credibility that have no acceptable workarounds generate a requirement for more detailed information in the appropriate category. This may include additional V&V and software testing and documentation, collecting additional test data, creating (and implementing) a CM plan, establishing new user support functions, or enhancing M&S functionality to meet the application requirements.

Activities required to generate this information are then included in the accreditation and V&V plans. The program manager (PM) wanting to use the M&S will have to provide the resources necessary to generate that information. Alternatively, if the PM cannot provide more funding or chooses not to do so, the PM can choose to use the M&S with the amount of evidence available and accept a higher level of risk. Some of the recommended activities for the example 6-DOF M&S are evident in Table 4; others were based on a list of recommended activities to support reducing an overall "moderate risk" assessment to "low risk" [2].

Conduct V&V and Other Credibility Activities

Once the accreditation and V&V plans have been developed and based on the preliminary risk

Table 2. Overview of M&S Requirements, Acceptability Criteria, and Metrics [3]

M&S REQUIREMENTS	ACCEPTABILITY CRITERIA	METRICS/MEASURES
General Capability		
The M&S shall support decision-making regarding the general intended uses.	Requirements documents and design are adequate for its intended use. Input data required for the M&S tool execution are available, and the degree of their validity can be established. The output parameters are appropriate for the intended use.	Documentation is available and complete. Input data requirements are documented. SME review determines that design is adequate, input data are credible, and output parameters appropriate.
Specific Capability/Accuracy		
The M&S shall have the functional, fidelity, and accuracy characteristics required to model the interrelationships affecting the specific intended uses.	The M&S models interrelationships between the following: (*Details of the requirements to be listed here.*)	SME reviews and approval of functionality and fidelity characteristics; SME assessment that the accuracy of outputs is acceptable considering the intended uses.
Software Accuracy		
Software shall be tested adequately to demonstrate its proper operation against the requirements identified.	Appropriate software test results and verification activities have been conducted; the software development environment is well structured and documented.	SME reviews and accepts software test and verification results and reviews and accepts artifacts of the software development process.
Data Accuracy		
Input and embedded data shall be adequate and appropriate for the application and documented.	Input and embedded parameters are appropriate for the intended use. Data sources are documented, appropriate, and authoritative. Data are sufficiently current for planned uses.	SME reviews and accepts input data and requirements, documentation, and data transformation verification.
Output Accuracy		
The outputs of the M&S shall be of sufficient fidelity and accuracy to support potential user requirements.	The dynamic behaviors are appropriate for the intended uses. Results are of appropriate fidelity for the intended use. Results compare favorably to other M&S (benchmarking), SME expectations (face validation), and/or available statistical analyses of comparisons with test data (results validation).	SME reviews, accepts, and determines output parameters important to the intended use.
Usability		
Processes and documentation shall be in place to ensure proper operation and appropriate interpretation and use of outputs.	Configuration management (CM) processes are sufficient and adequately documented and followed. Users are appropriately skilled and have the necessary training. User and analyst manuals and training are adequate to enable the user to properly execute the simulation and enable the user or analyst to properly understand outputs.	SME and user review and accept CM plans and artifacts, user training, experience, credentials, user and analyst manuals, and training materials.

Table 3. Partial Table for Air Vehicle 6-DOF [3]

M&S REQUIREMENTS	ACCEPTABILITY CRITERIA	METRICS/MEASURES
A. Atmospherics		
A1 Shall simulate turbulent environmental conditions using either the von Karman or the Dryden formulas.	A1 The M&S (incorporates either the von Karman or the Dryden form of turbulence models).	A1 SME review of comparisons between the wind output data (velocity, etc.) from the selected turbulence model and the expected turbulence form (von Karman or Dryden).
A2 Shall simulate the AV response to varying levels of wind gusts.	A2 The M&S incorporates an ability to induce varying strength gusts onto the AV.	A2 SME review comparing the wind output data (aircraft velocity, altitude, wind velocity, etc.) from the gust model, with expected gust model results.
B. Air Vehicle		
B1 Shall simulate mass properties of the AV.	B1.1 The M&S accepts a mass properties database.	B1.1 Documentation is available describing the process to incorporate a mass properties database file.
	B1.2 Mass property parameters output from the M&S agree with the expected output according to the database model.	B1.2.1 Verify that the mass property parameters output from the M&S agree with the expected output according to the database model. B1.2.2 SME reviews of documentation supporting the validation of the process used to create mass property database files (mass property model).

Table 4. M&S Characteristics Risk Assessment Criteria and Results Summary

CHARACTERISTIC	M&S CHARACTERISTICS IN THE RISK ASSESSMENT	INITIAL RISK ASSESSMENT FOR 6-DOF FLIGHT SIMULATION	
Capability (Criteria With Rating)			
Intended use and acceptability criteria	The general and specific intended use(s) of the M&S is/are clearly stated; the acceptability criteria and their metrics are clearly articulated (acceptability criteria and measures articulate how the requirements for the intended use[s] will be met).	Specific intended uses for the 6-DOF are clearly stated and approved by the program office.	Green
Conceptual model validation	The conceptual model (framework, algorithms, data sources, and assumptions) is documented and correctly and adequately describes the needs and requirements of the intended use.	Although conceptual model documentation is only available in multiple documents, the conceptual model has been reviewed in detail by SMEs for previous usage. The predecessor conceptual model has also been the subject of review, although documentation is largely unavailable.	Yellow

CHARACTERISTIC	M&S CHARACTERISTICS IN THE RISK ASSESSMENT	INITIAL RISK ASSESSMENT FOR 6-DOF FLIGHT SIMULATION	
Model fidelity (function- and entity-level decompositions)	The model's functions, entities, interfaces, data (framework, algorithms, data values, and assumptions), and environmental representation levels are documented and appropriate for the intended use.	Functionality and fidelity have been judged to be adequate by numerous SME reviews. Comparisons to actual flight test data have supported that SME consensus.	Green
Accuracy (Criteria With Rating)			
Design and implementation verification and validation	The algorithms and/or mathematical formulations are correct and valid. The premises for the application of the algorithms and/or mathematical formulations are correct, with no assumptions violated. Logical software implementation is correct and relatively error free.	Based on prior usage, the model design is sound and produces credible results. Detailed documentation of the algorithms should be developed.	Yellow
Input and embedded data	The simulation input and embedded data are credible and subject to review and revision.	Most input and embedded data are verified and validated via previous program activities and by comparisons with actual flight test data; however, they are not well documented. All current input data are via government-reviewed and accepted contract data requirements lists.	Yellow
System verification	The M&S architecture has been formally tested and/or reviewed and has been demonstrated to accurately represent the system simulated for which the SIUS, requirements, and acceptability criteria were articulated.	The model has been tested against actual prior system flight test data; the current model has not yet been formally tested.	Yellow
Output validation	The M&S responses have been compared with known or expected behavior from the subject it represents and has been demonstrated to be sufficiently accurate for the specific intended use(s).	The legacy performance model has been tested against actual flight test data and documentation provided for the aerodynamics and propulsion models. Similar validation for the current system with flight test data is planned but not yet executed.	Yellow
CM	The M&S and its components are under a sound CM process.	The development contractor uses Clear Case for configuration management of the software; however, there is no written CM plan.	Yellow
Usability (Criteria With Rating)			
Documentation	The M&S is well documented as to its capabilities, design and implementation, limitations and assumptions; the documentation is readily available, up-to-date, and complete.	Model documentation is informal; much of the available documentation is summarized in a draft Accreditation Support Package.	Yellow
User community	The M&S is designed and developed for the level of competency of the intended users. The users have access to documents such as user's manual, training manuals, and/or reference guides. User support is available from the M&S developer or proponent.	Program SMEs will run the simulation. There is a user manual; however, there is no analyst manual.	Green

assessment, the activities in those plans are implemented. Verification activities include determining and documenting whether the SIUS is correct and appropriate for current application; determining if defined capability, accuracy, and usability requirements are correct and complete for the SIUS; and determining if the capability, accuracy, and usability implementations are correct and appropriate per conceptual and design specifications and standards. Table 5 lists some example verification activities.

Validation activities include conducting and documenting data V&V checks, comparing simulation output/results to measured data (results validation) and/or an existing validated M&S output (benchmarking), and face validation (SME review). Sensitivity analysis can be a powerful tool supporting face validation reviews, developing requirements for validation test data, and analyzing M&S data against test results. Statistical analysis techniques play an important role in M&S output comparisons with test data. Some techniques that should be considered include Bayesian statistics, testing for intervals, and goodness of fit approaches like the Chi-Square, Kolmogorov-Smirnov (nonparametric), and Fisher's Combined Probability tests.

SME reviews of V&V data resulting from the 6-DOF V&V effort were conducted. The SMEs represented several interested organizations, including the unmanned aerial vehicle (UAV) program office, their support contractor (who was also running the 6-DOF as part of development), independent SMEs, and one or two representatives of operational

> *"Sensitivity analysis can be a powerful tool supporting face validation reviews, developing requirements for validation test data, and analyzing M&S data against test results.*

test and evaluation organizations who would have vested interests in the UAV program later in its development. The reviews resulted in some recommendations for further activities (which were planned for the next iteration of UAV development), considerable discussion of the technical merits of the 6-DOF and V&V results, and a consensus that the 6-DOF met the acceptability criteria for its intended use.

Table 5. Example Verification Techniques

TECHNIQUE	DESCRIPTION
Static testing	Uses source code-level static analysis tools and software quality checks (e.g., McCabe complexity index).
Dynamic testing	Exercises the software in its intended environment with a controlled set of inputs, hoping to replicate a predetermined set of results; unit testing, integration testing, regression testing, and white box testing.
Design verification	Tests software capability against measurable requirements.
Implementation verification	Verifies design requirements, code reviews, software error tracking, user documentation review, and sensitivity analysis.
Code reviews	May be line-by-line testing against design.
Black box testing	Confirms that the M&S implements the conceptual model, design, and requirements.

Update Risk Assessment and Iterate the Process as Needed

The VV&A process is iterated as necessary by updating the risk assessment, as tasks are completed in accordance with the V&V and accreditation plans. After all necessary iterations are completed, a "Final Risk Assessment" is developed and documented using the 10 M&S characteristics shown in Table 4. That final assessment determines the residual risk associated with applying the M&S to the intended uses after all V&V and other accreditation activities have been completed. That residual risk assessment, along with accompanying recommendations, is provided as supporting information to the accreditation authority. The final risk assessment for the 6-DOF M&S after all V&V activities were completed was "low."

Accreditation Assessment, Package, and Report

All metrics are ultimately adjudicated by SME review, making use of all information obtained and/or developed during the V&V process, which may include extensive comparisons to test data, benchmarking against other simulations, or comparing to a SME's perception of the real world. An accreditation decision ultimately relies on the accreditation agent/team leveraging personnel who know the subject matter to make a recommendation (based on the information available) as to whether the tool meets the requirements and can satisfy the intended uses and what should be done about it if it does not. Accreditation should be based on an objective comparison of the known credibility information with the credibility requirements, as shown in Figure 2. Based on that comparison, the accreditor can decide to accept the residual risk, require workarounds for risk areas or improvements to the M&S, or not accredit it at all.

The results of the V&V efforts and the accreditation recommendations are documented in a tailored MIL-STD-3022 [3] format. MIL-STD-3022 describes the standard formats for accreditation plans and reports and V&V plans and reports. The accreditation recommendation is documented in a letter from the M&S proponent (e.g., the M&S IPT lead) to the accreditation authority (the PM for program-related M&S uses), who is the final decision-maker and ultimate user of the M&S results as described in the SIUS. Based on all the work accomplished by the VV&A team and the 6-DOF developer, it was recommended that the 6-DOF be fully accredited to support its intended use.

> ***Accreditation should be based on an objective comparison of the known credibility information with the credibility requirements.***

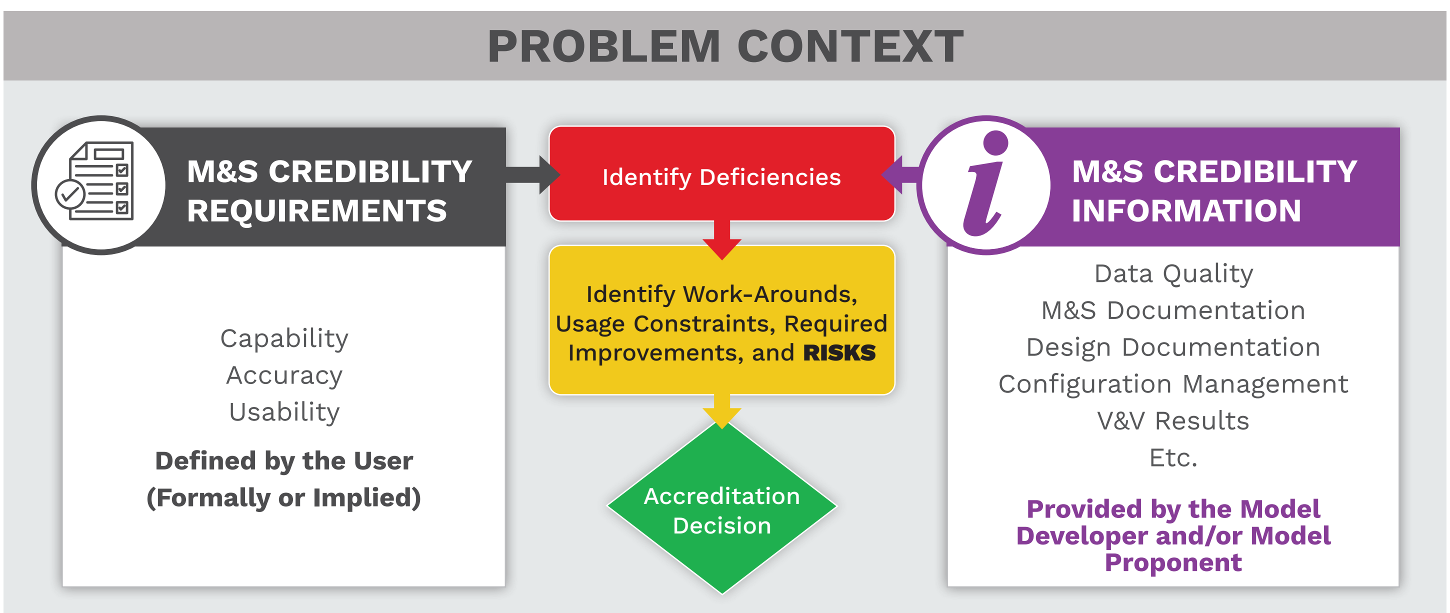


Figure 2. The Essence of Accreditation *(Source: NAVAIR VV&A Branch).*

CONCLUSIONS

This systematic VV&A process consists of determining, verifying, demonstrating, testing, and documenting whether the M&S requirements, acceptability criteria, and associated metrics and measures have been satisfied correctly. Because the M&S requirements are determined and defined from the SIUS, PMs, operational testers, and M&S users can have confidence in knowing whether the M&S has the credibility necessary to adequately support its intended use. This process applies to all M&S and test facilities, including live, virtual, and constructive simulations. What makes this process cost-effective is that any V&V activities are focused on requirements driven by the intended use; no V&V activities are conducted that do not directly support the requirements of the SIUS. ■

REFERENCES

[1] Law, A. M. *Simulation Modeling and Analysis*. 5th edition, McGraw Hill, ISBN 978-0-07-340132-4, p. 240, 2015.

[2] Naval Air Warfare Center. "Information Requirements In Support of Accreditation Volume II of the Accreditation Requirements Study Report." Computer Sciences Corporation for the Susceptibility Model Assessment and Range Test (SMART) Project, JTCG/AS-93-SM-20, Weapons Division, China Lake, CA, February 1994.

[3] U.S. DoD. *Department of Defense Standard Practice: Documentation of Verification, Validation, and Accreditation (VV&A) for Models and Simulations*. MIL-STD-3022, 28 January 2008.

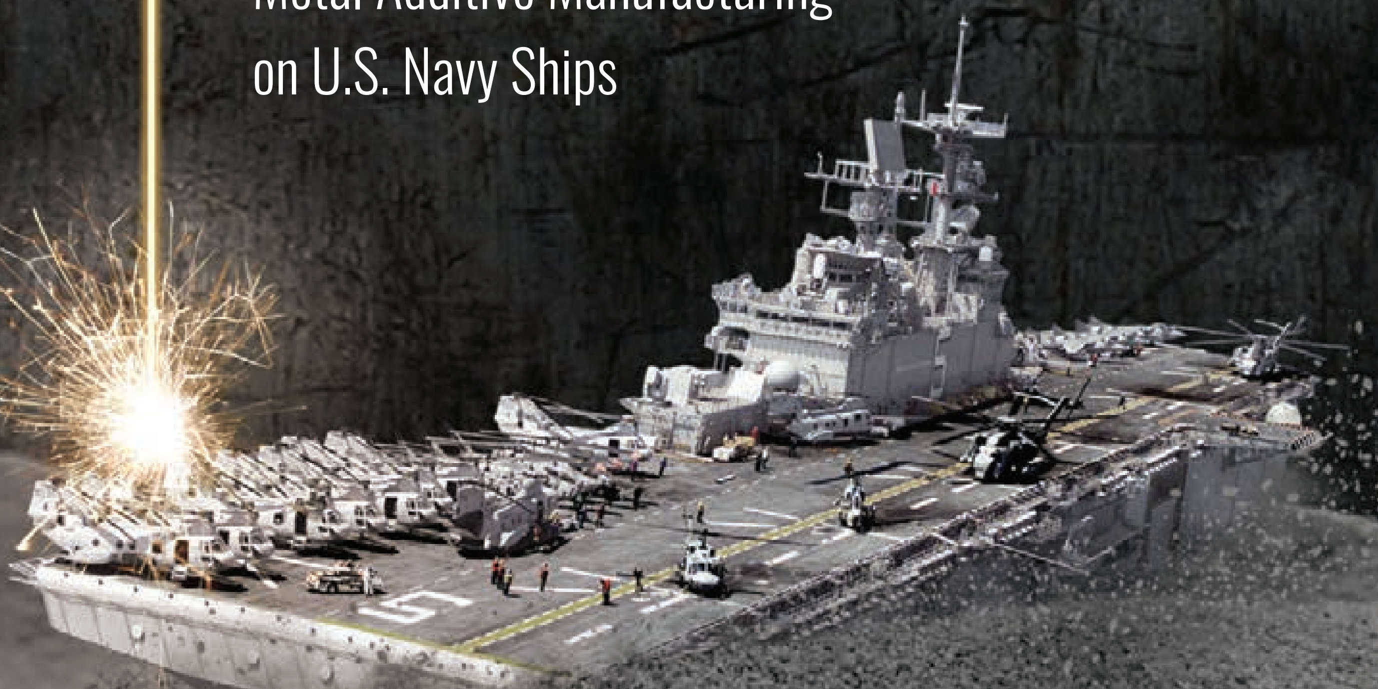
BIOGRAPHIES

DAVID H. HALL works for the SURVICE Engineering Company as the chief analyst for VV&A and analysis support services under contract to the NAWCAD. Mr. Hall holds B.S. and M.A. degrees in mathematics from California State University at Long Beach.

DAVID J. TURNER is the Patuxent River area operation manager for the SURVICE Engineering Company, supporting a wide variety of M&S VV&A programs at NAWCAD. Mr. Turner holds a B.S. degree in aerospace engineering from Pennsylvania State University.

REVOLUTIONIZING NAVAL LOGISTICS:

The Challenges and Prospect of Metal Additive Manufacturing on U.S. Navy Ships



BY MATTHEW SEIDEL

(PHOTO SOURCE: U.S. NAVY, CANVA)

THE CURRENT STATE OF METAL ADDITIVE MANUFACTURING (AM)

In 2022, the U.S. Navy installed the first-ever permanent metal AM machine aboard a U.S. naval vessel. This technology is projected to be groundbreaking by reducing resupply logistics and diminishing obsolescence: "For the Navy, the greatest immediate potential is in the less-exotic field of logistics" [1]. However, getting a contraption that can create almost any replacement part onto an active-duty warship has had its challenges and obstacles. This article will focus on metal AM and detail the development of concepts from the early 2000s until now, along with the history of challenges in the field.

AM has experienced a renaissance in the last decade. Formerly only available for use by large companies with extensive capital expenditure budgets, modern AM has significantly decreased in price, branching out into applications for industry and hobbyists alike. Naval planners took note of this technology and followed suit with other government bodies in proposing shipboard augmentation of supply departments with AM. This would reduce the need to carry large stocks of premanufactured replacement parts, diminish delivery costs and timelines for spare components, and free up coveted space on dense vessels.

> ***AM has experienced a renaissance in the last decade.***

CHALLENGES OF SHIPBORNE INTEGRATION

As integrating AM into the Navy's logistical operations quickly became a priority, challenges rose to the forefront. AM machines are typically set up in a laboratory or industrial setting that offer ample space for a large machine, a stable building, and storage for many different printing materials. If the Navy were to install this capability into ships, they would need to address the harsh environments and limited resources of being on the open ocean. For example, a popular industrial version of metal AM is selective laser sintering, which involves powdered metal being melted layer by layer as new powder is wiped on top of the solidified layer. This technique creates high-quality parts that can equal or surpass traditionally manufactured parts created from sources like titanium, Inconel, and steels. However, the powdered metal, particularly the titanium variety, can become explosive with improper handling due to its high surface area-to-volume ratio interacting with oxygen in the air.

In an industrial setting, this is managed by requiring special safety measures and trained personnel to handle the raw materials. In a shipborne setting, the roll of the ocean could disrupt this powder, causing hazardous spills. Certain metal AM technologies may be optimized for a lab setting; however, they could prove disastrous in a shipborne setting. This potential detriment could easily outweigh the benefits of a vertically integrated AM machine. Therefore, determining what kind of raw material to use was an important consideration.

Any AM system requires consistent electrical power for quality part manufacturing, and shipborne power generation can experience fluctuations. Most ships rely on auxiliary engines, generators, and/or shore power to provide electricity, all of which are susceptible to fuel supply issues and power transmission issues. Power fluctuation can also occur when other heavy demand systems come online, stressing the grid and giving uneven electrical distribution. Scheduled maintenance and repairs may even become an issue, as this can occur on one side of the ship and affect systems on the other side. A power supply disruption, even for a few milliseconds, could immediately affect the ongoing printing process and potentially lead to incomplete or failed prints. Onboard AM machines would need to have their own semi-isolated grid or specific equipment to ensure there are no power issues during the critical printing process.

Other environmental issues to consider on a ship would be the constant motion caused by waves and tides. This could change the powder distribution or print head angle and result in poor print quality and wasted raw materials. Vibrations from ship engines could have deleterious effects on the precision of AM machines.

The notorious salt fog and salty air environment of an ocean-going ship could wreak havoc on an AM machine designed for the sterile laboratory setting. A shipborne AM machine would have to be protected from all these conditions to bring the benefits anticipated by naval logistics. Due to the untested nature of having AM onboard a ship, there was an additional X factor consideration. And beyond the known hazards were the unknown risks that scientists could anticipate or predict.

The idea of having an AM machine accessible on every ship in the United States' fleet also presents its own challenges. Many AM machines require an internal atmosphere of inert gases which would need to be resupplied occasionally. Any machine that needs a special gas atmosphere must consider off-gassing, proper containment, leak prevention, and proper personal protective equipment (PPE). A typical ship's interior is a self-contained atmosphere. Without air circulating, harmful gases can easily fill a watertight room. Then the feedstock itself is not available at every resupply station, as it can only be produced through specialized manufacturing processes. Training and PPE would also need to be addressed because of the hazardous equipment and materials used in the printer.

> *The notorious salt fog and salty air environment of an ocean-going ship could wreak havoc on an AM machine designed for the sterile laboratory setting.*

THE FIRST NAVY SHIPS TO RECEIVE METAL AM

In the latter half of 2022, the Navy's Wasp-class amphibious assault ship USS *Bataan* (Figure 1) received the first-of-its-kind, permanent metal AM machine. Based in Norfolk, VA, the nearly 850-ft "Harrier Carrier" supports Navy and Marine Corps teams and has the capabilities to print on-demand replacement parts. The Wasp-class ships were chosen to receive the first metal AM machines due to the role they play in the U.S. military. These ships house both Navy and Marine forces and materiel, which allows a wider avenue of research and exposure on parts where form and fit are key across several different platforms for both branches. During its most recent outfitting in November 2022, the USS *Bataan* was equipped with a Phillips additive

Figure 1. USS *Bataan (Source: Adlughmin [3]).*

hybrid powered by Haas [2], which consists of a combination of a Haas TM-1 computer numerical control (CNC) mill and a Meltio laser metal wire deposition head. The Haas TM-1 platform has been a proven platform on ships before and therefore provides minimal new variables for the Naval Sea Systems Command (NAVSEA) to overcome when testing the AM features of this system.

The advantage of this system, according to Phillips Corporation [4], is that it is a traditional CNC subtractive manufacturing system with an added AM capability. The additive features of this machine are like welding, in that a metal wire like 316L stainless steel is fed into a print head where focused lasers melt it to a rough shape. The process begins with a metal powder material being fed into the system, where it is melted and deposited layer by layer onto a substrate using a high-power laser to a precision of ±0.010 inches. The subtractive feature of the hybrid system then comes into play, the HAAS CNC machine mill refines the rough shape out using traditional CNC milling, and the result is a highly accurate part. The overall hybrid AM system generates significantly less waste compared to traditional manufacturing.

316L stainless steel, used in the Phillips additive hybrid machine [4], is a popular material choice for marine environments due to its excellent corrosion-resistance properties. The "L" in 316L stands for low carbon, meaning it has a reduced carbon content compared to other grades of stainless steel. This lower carbon content helps prevent sensitization, a process where carbon combines with chromium to form chromium carbide, which can cause the material to become susceptible to corrosion. Additionally, 316L stainless steel has good strength and toughness properties, as well as excellent weldability and formability, which makes it easy to fabricate into complex shapes and structures.

In terms of applying this hybrid AM machine, sailors will have the capability to print over 300 NAVSEA-developed AM technical data packages (TDPs) on demand. In addition, parts can be repaired by a method similar to cold spray, i.e., adding back material to a broken section of an existing part and then machining back excess material to repair the broken part.

In July 2022, another Wasp-class ship, the USS *Essex*, received a different style of AM technology [5]. In Pearl Harbor, HI, the ElemX made by Xerox was lifted onto the USS *Essex* via a Conex box into the cargo bay of the ship. (This machine will be permanently housed inside the metal container, protecting it from the environment.) The "mini factory in a Conex box" was previously installed at the Naval Postgraduate School in Monterey, CA, in 2020. For two years, testing was conducted there on this unique solution to assess its capabilities for incorporation into naval vessels. With this prior research well documented, the results of the laboratory setting of ElemX, in which control parts were printed, will be compared with results gathered from the deployed USS *Essex* [5].

The ElemX claims to be a user-friendly metal AM machine—no hazardous metal powders and no need for extensive facility modifications or PPE. It uses standard aluminum wire melted into a recyclable powder support. The printer utilizes a proprietary "liquid metal" technology; "unlike alternative AM technologies, there are no metal powders used with ElemX and no need for PPE or other considerable safety measures. Engineered to bring simplicity to the supply chain process, ElemX is said to be the ideal option for spares, repairs, and low-volume production parts" [6].

Liquid metal printing utilizes the same concepts that Xerox used decades ago

> ***Liquid metal printing utilizes the same concepts that Xerox used decades ago in inkjet printing but substitutes ink with liquid metal and allows printing in a third dimension.***

in inkjet printing but substitutes ink with liquid metal and allows printing in a third dimension [7]. Aluminum wire is fed from a spool onto a "hopper" surrounded by a copper wire with a pulsed voltage, which melts the metal wire into a liquid and deposits it to a heated "substrate" where it is solidified. The heated build plate also requires a noble gas shroud environment like welding and protects the molten pool of metal against the elements in the atmosphere. Like in other three-dimensional printing styles and technologies, liquid metal printing requires its own unique set of parameters that must be optimized and tested for size of ejected drop, droplet ejection, and thermal diffusion from the droplet.

The Xerox ElemX uses A356/4008 aluminum alloy wire, which has been used in a variety of marine applications, such as boat hulls, propellers, and other marine structures. Its combination of excellent corrosion resistance and high strength-to-weight ratio makes it an ideal material choice for marine environments. It is a high-silicon alloy that contains 7% to 9% silicon, which gives it superior corrosion resistance compared to other aluminum alloys. This high-silicon content creates a dense, protective oxide layer that helps to prevent corrosion in marine environments, where exposure to saltwater and other corrosive agents is common.

SHORESIDE TESTING AND COLLABORATION

The objective of these printers onboard the USS *Bataan* and *Essex* is for current testing. As discussed earlier, there are many factors theorized to affect print quality. But there is also a concern that there are unknown factors yet to be fully realized that need to be addressed and solved before full operation and expansion to other ships can occur. Speaking with Professor Ibrahim Gunduz at the Naval Postgraduate School (NPS), his group aims to assist finding those unknown factors in a joint service effort (NAVSEA, U.S. Coast Guard, Marines, and Army) [8].

NPS will also provide recommendations for equipment, engineering standards, academic studies, and testing for AM to the group and oversee the land-based testing for the ElemX machine currently on the USS *Essex*. This involves installing the printer on their premises in California and printing a series of laboratory test prints and measuring every aspect of the result, including mechanical properties, correlating environmental data, and TDPs. These first test prints will then be directly compared to identical parts made while the same machine is deployed at sea. The twin tests conducted at sea will use an array of onboard sensors, such as pressure, humidity, gyroscopes, etc. This instrumentation will determine and monitor atmospheric and motion parameters. NPS will then feed the results of the experiments back to the joint AM group. When addressing the foreseeable issues of printing with metal on a ship, these two printers will utilize different AM technologies to address some of the issues in different ways. According to NPS, this was a choice made consciously and chosen so that they complement each other [8].

It is also speculated that using artificial intelligence (AI) or machine learning (ML) aboard these ships can increase the reliability of AM [9]. One of the key advantages of employing AI and ML in AM is the ability to enable real-time quality control. Instead of relying solely on post-processing inspections, which can be time-consuming and costly, the AI system can provide immediate feedback during the printing process. This allows early identification and rectification of issues, minimizing the number of failed prints and reducing material waste. Using sensors, AI algorithms can be trained to identify defects and anomalies in the AM process, enabling real-time quality control, minimizing

> ***One of the key advantages of employing AI and ML in AM is the ability to enable real-time quality control.***

the need for post-processing inspections, and reducing the amount of time wasted by identifying or correcting failing prints before they finish.

The Navy recognizes that in the long term, neither of these printers can print large items like a torpedo or a dinghy, but it does provide a great solution for small, durable parts in relatively low quantities. For example, the idea of printing custom-built drones has been of great interest to naval planners from the beginning of this effort. This would keep service men and women out of harm's way while being able to have a drone built for a certain mission [9]. Data files can be sent via satellite to the USS *Essex* to quickly build a custom frame.

Back in 2015, while the USS *Essex* had a polymer AM machine installed, a test quadcopter drone frame was printed and fitted with a transmitter and a camera [3]. Its mission was to fly over ships to help stop piracy and drug smuggling. The evolution from polymer to metal could enhance drones' capabilities, larger payload, longer flight time, and the ability to fly in harsher conditions. The addition of these printers can allow warfighters to carry less cargo for every perceivable drone mission and only carry raw printing materials and electronics; the same would apply to another mission where parts could be printed, depending on what is needed at that time.

SUCCESS STORIES ONBOARD METAL AM OPERATION AT SEA

In August 2023, the crew of the USS *Bataan* successfully used the Phillips Additive Hybrid System to create and replace a sprayer plate for a de-ballast air compressor while at sea [10]. The metal sprayer plate was essential for forcing pressurized air through saltwater tanks to discharge accumulated saltwater, a process used to lower the ship's draft for amphibious operations.

The repair effort was led by Machinery Repairman First Class Mike Hover, who created a computer-aided design (CAD) model of the sprayer plate using a functional one from another system as a reference [10]. NAVSEA's "Apollo Lab" provided engineering support and training, refining the CAD file. Mechanical engineer Bryan Kessel at the Naval Surface Warfare Center, Carderock Division worked on software instructions for the metal AM machine. These instructions were securely transferred back to the ship, where the sprayer plate was produced and installed. The replacement part saved time that would have been spent obtaining a replacement assembly. This achievement, completed in just five days, marked the first time the ship's installed metal AM machine was used under such conditions for an actual repair.

CONCLUSIONS

The journey to implement AM on ships has presented challenges such as addressing environmental factors, ensuring safety measures, and optimizing print quality. The U.S. Department of Defense bodies like NAVSEA, academic institutions like NPS, and industry leaders in AM technology are all actively contributing to our understanding of and providing solutions for these challenges. The promise of permanent metal AM machines on vessels like the USS *Bataan* and *Essex* opens new possibilities for reducing resupply logistics, mitigating obsolescence, and enabling on-demand production of replacement parts. Success stories like the de-ballast air compressor repair give promise to the Navy's efforts to leverage AM technology to enhance readiness and self-sufficiency in maintaining ships and weapons systems in challenging operational environments. This successful application of AM technology demonstrates the "tip of the iceberg" for capabilities that can be achieved. ■

REFERENCES

[1] Cheney-Peters, L. S. "Print Me a Cruiser!" U.S. Naval Institute, https://www.usni.org/magazines/proceedings/2013/april/print-me-cruiser, accessed on 28 June 2023.

[2] Listek, V. "U.S. Navy Installs Meltio Hybrid Metal 3D Printer to Reduce Repair Times." 3DPrint.com, https://3dprint.com/297004/us-navy-installs-meltio-hybrid-metal-3d-printer-to-reduce-repair-times/#:~:text=In%202022%2C%20the%20USS%20

Essex,printer%20on%20November%203%2C%2 -02022, accessed on 28 June 2023.

[3] Adlughmin. "U.S. Navy Is 3D Printing Custom Drones Onboard the USS *Essex*." 3DPrint.com, https://3dprint.com/85654/us-navy-3d-printed-drones/, accessed on 28 June 2023.

[4] Dhananjay. "U.S. Navy Installs on Board the First Phillips Additive Hybrid Metal 3D Printing Solution Powered by Meltio and Haas." Phillips Corporation, https://www.phillipscorp.com/us-navy-installs-on-board-the-first-phillips-additive-hybrid-metal-3d-printing-solution-powered-by-meltio-and-haas, accessed on 28 June 2023.

[5] Verger, R. "A Navy Ship Got a Giant Liquid-Metal 3D Printer." *Popular Science*, https://www.popsci.com/technology/navy-ship-gets-large-metal-printer, accessed on 28 June 2023.

[6] *Metal AM Magazine*. "Vertex to Offer Contract Manufacturing With Xerox ElemX Liquid Metal." https://www.metal-am.com/vertex-to-offer-contract-manufacturing-with-xerox-elemx-liquid-metal, accessed on 28 June 2023.

[7] Ruggles, A. "Liquid Metal 3D Printing." *FLOW*, https://www.flow3d.com/liquid-metal-3d-printing/#:~:text=Ejected%20droplets%20travel%20to%20a,deposition%20of%20the%20incident%20droplets, accessed on 28 June 2023.

[8] Curran, C., I. Gunduz, and CAPTs J. Gray and G. Hobson. Personal communication, 7 March 2023.

[9] Isi, C. "AI-Assisted 3D Printing: Insights on Emerging Trends and Technologies." Protolabs Network, https://www.hubs.com/blog/ai-assisted-3d-printing, accessed on 28 June 2023.

[10] *Metal AM Magazine*. "NAVSEA Improves Readiness of USS *Bataan* With On-Board Metal Additive Manufacturing." https://www.metal-am.com/navsea-improves-readiness-of-uss-bataan-with-on-board-metal-additive-manufacturing, 29 August 2023.

BIOGRAPHY

MATTHEW SEIDEL is a mechanical engineer working with Innvometric to develop inspection and quality control software. He is an expert in metrology, AM, and shipborne systems. He has worked as a design engineer for the Navy, a dimensional metrologist, and a consultant ship's engineer for private yachts. Mr. Seidel holds a bachelor's degree in mechanical engineering from the South Dakota School of Mines and Technology.

A SURROGATE MODEL TO QUANTIFY UNCERTAINTY

IN THERMAL PROTECTION SYSTEMS FOR HYPERSONIC WEAPONS

BY ERIC A. WALKER, JASON SUN, AND JAMES CHEN

(PHOTO SOURCE: DARPA)

SUMMARY

Modeling and simulation are key for the iterative development of thermal protection systems (TPS's) for hypersonic weapons. In this work, the temperature-dependent flexural strength (FS) of α-SiC ceramic is predicted given Young's modulus, Poisson's ratio, and temperature. An artificial neural network (ANN) surrogate model is created to retain property-performance prediction while increasing computation speed. The ANN computes many times faster (less than a second vs. tens of minutes) than a finite-element model (FEM). An uncertainty quantification (UQ) is necessary because property inputs vary due to defects in manufacturing processes. Here, the ANN surrogate model provides the necessary computational speedup to perform a UQ. A temperature-dependent UQ is demonstrated using the ANN. This work demonstrates that a machine-learning surrogate model is a useful replacement for a physics-based FEM for predicting the temperature-dependent FS of a TPS with UQ.

INTRODUCTION

The United States is currently developing hypersonic weapons [1]. Other countries have been ahead in deploying hypersonic weapons since 2016 [2]. A limiting factor in deploying hypersonics is their TPS's [3]. Hypersonic flight increases the heat resilience necessary of the weapons' TPS's. Without a TPS, critical components of the weapon will likely melt. To engineer a TPS for a hypersonic weapon, physics-based modeling is useful; however, it is slow. Furthermore, due to manufacturing defects, material properties are uncertain and must be considered, placing them outside the computational reach of physics-based simulations. UQ requires many samples of a model to effectively map the uncertainty in material properties to uncertainty in the TPS material's performance. Here, a fast surrogate model is a substitute for the full physics-based model. This model consists of an ANN with property inputs of Young's modulus and Poisson's ratio to FS, a performance measure, as the output.

The most prevalent UQ model of α-SiC is from the 1990s and uses purely data-driven parameters [4]. This model lacks physics, can potentially lead to unrealistic conclusions, and is a fitted logistic function. Silicon carbide is a temperature-resistant lightweight material that is used in hypersonic weapon TPS's. UQ can coordinate with decision makers in a manufacturing supply chain, where the risk of a material's failure on its final deployed state is a critical factor influencing decisions [5, 6]. A known problem with ceramics manufacturing is variations in microstructure [7] and surface defects [8, 9] from manufacturing processes. Hypersonic weapons are also exposed to particulates and acoustic waves during flight [10]. Therefore, a need for UQ in ceramics manufacturing is undeniable [11–13]. Recently, an experimentally driven approach has been taken in UQ of FS [14]; however, more developments in a generalizable model would be greatly beneficial.

The computationally light replacement model for the physics-based model is a surrogate model. Surrogate models stand in for the physics-based model and retain the nonlinear mapping from inputs to outputs. The surrogate model always works in tandem with the physics-based model by training on it. The physics-based model is solved with high-performance computing (HPC) and provides quality training data to the surrogate model. While successful in making predictions, physics-based models are computationally expensive [15, 16]. This computational expense limits their usefulness in UQ studies [17]. Even models that compute rapidly with HPC [18] do not approach the

> ***The computationally light replacement model for the physics-based model is a surrogate model.***

speed needed for a surrogate model for UQ. Should quantum computing become ubiquitous, large systems of linear equations that are the kernel of physics-based simulation could provide a powerful solution, but that is a distant horizon [19, 20]. Using the surrogate model here, the uncertainty in Young's modulus and Poisson's ratio may be adjusted and the model provides near-instant feedback on the resulting changes in FS uncertainty. The key to this rapid UQ will be the machine-learning (ML) model.

The ML model applied in the ANN converts an input vector, including material properties and temperature, to FS at a wide range of temperatures. In the next section, the ANN theory is summarized and discussed. The implementation of ANN with the backpropagation algorithm is outlined in the ANN Implementation section, where two optimization algorithms are compared. In the ANN Training section, the accuracy of the ANN is tested by randomly removing data from the training set and predicting on those removed data. The number of data records removed is incrementally increased to the breakdown's limit. A temperature-dependent forward UQ from Young's modulus and Poisson's ratio to FS is then conducted using the ANN surrogate model. The interpolation among temperatures not computed by the thermomechanical fracture model is shown to operate properly. The temperature dependence is important for demonstrating the model's capabilities for hypersonic weapons' TPS's. Finally, the outlook for these weapons' ANN and manufacturing is discussed in the Conclusions section.

ANN THEORY

An ANN consists of layers of neurons with connections between them. A neuron stores a value and receives and sends signals (which are essentially values) or numbers. The connections map the output signals from one layer of neurons to the next layer of neurons, where the signals become inputs. Weights on each connection are trained. Each neuron also has a constant bias that is trained. The ANN in this work is supervised by FSs as the labels. The ANN in FS modeling is broken down next.

The general equation for the value of given input χ_j to a given neuron is

$$\chi_j = \sum_i y_i w_{ji}, \qquad (1)$$

where y_i are inputs from neurons from a previous layer and χ_j is the sum of the signals to neuron index j closer to the outputs [21]. w_{ji} proceeds from left neuron index i nearer to inputs to right neuron j nearer to outputs. A bias for a given neuron is added with an additional input value of 1, and its weight is then treated the same as other weights. Next, the cost function, or error E, is defined with the difference of the labeled data point and the output from the ANN as

$$E = \frac{1}{2}(y_{FS} - d_{FS})^2, \qquad (2)$$

where d_{FS} is the labeled data point and y_{FS} is the output from the ANN, given the inputs and weight values [21]. To minimize the cost function, E, the gradient or sensitivity of the cost E to each of the weights is sought. The naive way to obtain the gradient would be to perturb each weight one at a time by reevaluating the ANN at every step. Obtaining the gradient would be a computationally expensive process. Backpropagation is a computationally efficient algorithm to provide the gradient.

In the backpropagation algorithm, the ANN is only run forward one time. All the χ's and y's are found from the forward pass. Expressions for the gradient are written using derivatives and the chain rule from calculus. These expressions are orders of magnitude less computationally burdensome than evaluating the entire ANN again. Starting with the output layer, the partial derivative is taken for the output signal, y_{FS}, as follows [21]:

$$\frac{\partial E}{\partial y_{FS}} = y_{FS} - d_{FS}. \qquad (3)$$

The derivative of the cost regarding the total of the input signals to the output neuron y_{FS} is taken as

$$\frac{\partial E}{\partial \chi_{FS}} = \frac{\partial E}{\partial y_{FS}} \frac{d y_{FS}}{d \chi_{FS}}. \qquad (4)$$

The term $\partial E/\partial y_{FS}$ is known from equation 3. χ_{FS} is the total input to the neurons in the final layer. The term $dy_{FS}/d\chi_{FS}$ is the derivative of the output signal for the total inputs. The rectified linear unit activation function is used in this work to prevent obtaining negative output [22]. The rectified linear unit activation is as follows [21]:

$$y_j = max(0, \chi_j). \tag{5}$$

For numbers greater than 0, $dy_{FS}/d\chi_{FS} = 1$. Thus, equation 4 becomes

$$\frac{\partial E}{\partial \chi_{FS}} = y_{FS} - d_{FS}. \tag{6}$$

From here, the gradient regarding the weights is needed. The derivative is written for the cost function for weights w_{FSi} on the signals leading to the final layer as follows [21]:

$$\frac{\partial E}{\partial w_{FSi}} = \frac{\partial E}{\partial \chi_{FS}} \frac{\partial \chi_{FS}}{\partial w_{FSi}}. \tag{7}$$

The expression $\partial E/\partial \chi_{FS}$ is known from equation 6. The expression $\partial \chi_{FS}/\partial w_{FSi}$ equals y_i, the signals from the last layer before the output layer (see equation 1). The process backpropagates to obtain the gradient for all the weights. Figure 1 shows the last layer of the neural network where the backpropagation algorithm begins. Figure 2 displays the input parameters and output prediction of the neural network, which has two layers between inputs and outputs and is 50 or 100 neurons wide.

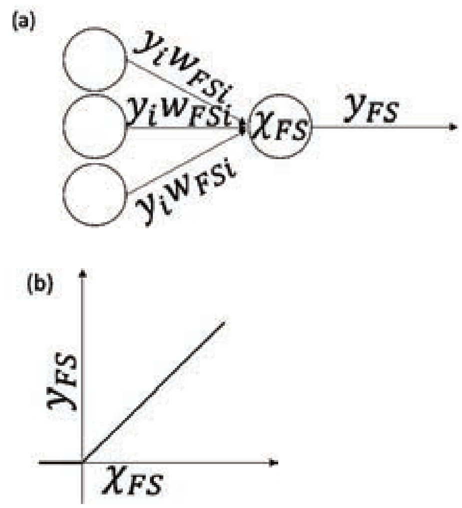


Figure 1. ANN Near the FS Output: (a) End of Neural Network and (b) Rectified Linear Unit Signal *(Source: E. Walker, J. Sun, and J. Chen).*

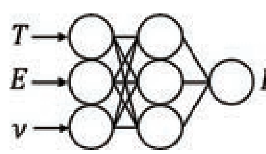


Figure 2. Inputs and Outputs of the ANN Surrogate Model *(Source: E. Walker, J. Sun, and J. Chen).*

ANN IMPLEMENTATION

Key computational details about ANN include the number of hidden layers between inputs and outputs, the width or number of neurons in those hidden layers, the activation function, and the algorithm choice for optimizing weights during training. The ANN in this work uses widths of hidden layers of about 100 neurons. After several tests, two hidden layers are deemed necessary and used. Dropping to one hidden layer decreases the R^2 on fitting back to the training data to 0.994. A near-perfect R^2 of 1 is expected in this scenario of predicting the same data used for training. Adding a second hidden layer does not increase the computation time by a noticeable amount. The weights are trained by a (limited memory) Broyden-Fletcher-Goldfarb-Shanno (BFGS) [23] algorithm. The maximum number of iterations to optimize the weights is set to 500. Poisson's ratio is on a smaller scale than Young's modulus and temperature. The inputs are rescaled to improve the training performance.

BFGS is a quasi-Newton method to minimize the cost function. It includes curvature information via a Hessian matrix in its search. The Hessian contains second derivatives of the cost function regarding the weights. A starting point for BFGS is Newton's method [23]:

$$x^{k+1} = x^k - \alpha^k [\mathbf{H}^k]^{-1} g^k, \tag{8}$$

where k is the iteration index, x is the vector of weights, and $\mathbf{H}$ is the Hessian. The inverse of the Hessian matrix is taken in Newton's method. However, since matrix inversion is computationally expensive, an approximation to the inverse Hessian matrix $[\mathbf{H}]^{-1}$ is computed instead. The gradient supplied by backpropagation is g^k. A scalar number α controls step size and is optimized at each iteration in BFGS. The BFGS updated equations are listed next. For x^0 in the initial guess state, a starting approximation matrix of the inverse

> ***BFGS is a quasi-Newton method to minimize the cost function.***

Jacobian, H^0, is guessed. A new compact variable s is defined by the following [23]:

$$s^k = -H^k g^k. \tag{9}$$

Then, the cost function E from equation 2 is minimized by introducing a scalar α^k,

$$\min_{\alpha^k} E(\alpha_k) = \frac{1}{2}(y_{FS}(x^k - \alpha^k s^k) - d_{FS})^2. \tag{10}$$

where x^k is the current vector of weights and biases. s^k is set by equation 9. Therefore, in equation 10, the only variability in the cost function $E(\alpha^k)$ is due to α^k. Methods exist to minimize $E(\alpha^k)$ by changing α^k [23]. Next, another compacted variable σ is introduced as

$$\sigma^k = \alpha^k s^k. \tag{11}$$

The weights and biases are updated as

$$x^{k+1} = x^k + \sigma^k. \tag{12}$$

The approximate inverse Hessian matrix is updated as

$$\mathbf{H}^{k+1} = \mathbf{H}^k + \mathbf{D}^k, \tag{13}$$

where the key update matrix **D** is obtained by satisfying another equation:

$$y^k = q^{k+1} - q^k, \tag{14}$$

where y is another compacted variable from other known variables,

$$y^k = g^{k+1} - g^k. \tag{15}$$

Another solver is a stochastic gradient-based approach known as "Adam" [24]. Parameters are updated with

$$x^{k+1} = x^k - \alpha \frac{\hat{m}_{k+1}}{(\sqrt{\hat{v}_{k+1}} + \epsilon)}. \tag{16}$$

The difference thus far from BFGS is that the Hessian matrix is not used in the update but rather parameters that are functions of g, $\hat{m}$, and $\hat{v}$, with ϵ as a constant hyperparameter. Adam is suited for large amounts of noisy training data [24]. As a test, BFGS and Adam were used to train two otherwise identical neural networks. The two trained neural networks were predicted back on the training data. BFGS was more accurate than Adam, as measured by R^2.

The ANN after training by BFGS will be run many times to predict FS, the output of choice in this study. This forward UQ maps the uncertainty in the three inputs to the FS. The inputs are sampled according to their probability density functions (pdf's).

ANN TRAINING

The training data set used in this study comes from a three-way coupled thermomechanical fracture model solved by the FEM. There are extensive efforts of developing physics-based simulations which can be used to study the bending, buckling, vibration, and stress intensity of various shapes of materials [18, 25–28]. Various attempts have been made to model the fracture behavior of brittle materials like ceramics [29–32]. The model generates the training data set on the FS of α-SiC over a wide range of temperatures, which is particularly helpful for applications like thermoprotection systems [33].

The three-way coupling model consists of modules, including elastic mechanics, phase field for damage, and heat conduction. The phase field uses an auxiliary scalar field to model material damage. This field is used to capture both the history of strain energy and the weakening effect on material properties as damage accumulates. The evolution of damage is governed by the Allen-Cahn equation [29, 30, 32]. The material properties are obtained from experimental data by Munro [4]. The simulation setup is based on a standard four-point bending test specified in ASTM C1161-18 [34]. The point of fracture is determined according to the Griffith theory [35]. The cell size is controlled to be small enough to capture the fracture surface [29, 30]. FS is defined as the maximum tensile stress that the specimen sustains before cracking.

The temperatures of the training data set from the numerical simulation are 400, 600, 800, 1000, 1200, and 1400 °C. The mean Young's modulus is as follows [4]:

$$E(GPa) = 415 - 0.023T(°C), \tag{17}$$

and the mean Poisson's ratio is

$$\nu = 0.16 - 2.62 \times 10^{-6} T(°C). \tag{18}$$

Together with the mean, four combinations of E±3% and ν±25% at each temperature fill out the parameter space as suggested by the experimental data [4]. There are five training points at each of the six temperatures, totaling 30 training points. The five data points are the mean of E and v, the upper bound of E with the upper bound of v, the upper bound of E with the lower bound of v, the lower bound of E with the upper bound of v, and the lower bound of E with the lower bound of v. For Young's modulus and Poisson's ratio, respectively, 3% and 25% are taken as the three standard deviations distance or 99% confidence interval when performing the forward UQ later. Each training data point consists of the three input parameters—Young's modulus, a Poisson's ratio, and a temperature, resulting in FS as the output. A similar approach was taken by Nagaraju et al. [36]. However, this work used an ANN and included adding temperature variation, 400–1400 °C, to emphasize the temperature effect.

After the ANN is trained by the 30 data points, a rigorous cross-validation test of the ANN accuracy is conducted and summarized in Figure 3 [37], where the FEM [33] provides 30 FS's from inputs to complete the training data set. Of the data set of 30 points, n data are selected randomly to move from training to testing. For instance, when n = 3, the training data set size is 27 and the testing data set size is 3. The ANN, after being trained on 27 data, predicts on 3 data. The R^2 tends to decrease with shrinking training data set size because less and less of the parameter space is covered during training. In Figure 4, five trials are conducted at each n. When n = 7, the first negative R^2 appeared. Therefore, the ANN is reliable with an n smaller than 7.

The process employed for validating the ANN is known as cross-validation [37]. The loss in reliability may be attributed to declining dimensionality of the training data or the variety in Young's modulus, Poisson's ratio, and temperature.

Now that the ANN is trained, one million samples are used for a forward UQ of α-SiC at 1400 °C (Figure 5) as an example. Both Young's modulus and Poisson's ratio are modeled as Gaussian. Their mean and standard deviation are also displayed in Table 1. For future studies, a Bayesian update could be used to change the means and standard deviations of Young's

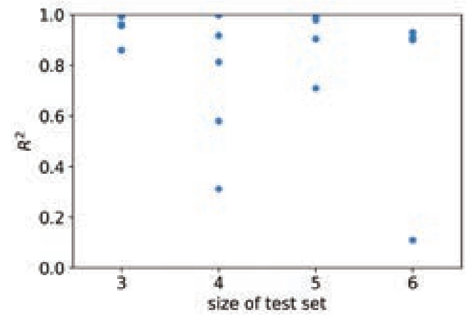


Figure 4. Cross-Validation of the ANN *(Source: E. Walker, J. Sun, and J. Chen).*

modulus and Poisson's ratio as more experimental data become available [38].

Figure 5 shows the UQ of FS of α-SiC at 1400 °C as an example. The mean of predicted FS is 417.03 MPa, compared with 416.84 MPa computed by the physics-based model. The ANN therefore achieves consistency with the physics-based model. With this tool, a manufacturer can rely on the ANN

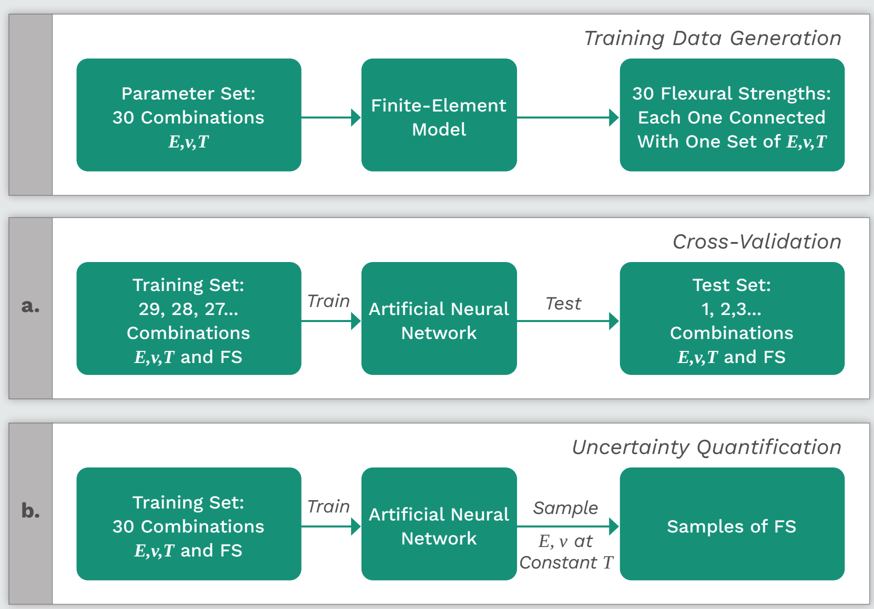


Figure 3. The FEM Training Data Set: (a) Cross-Validation for the ANN Reliability and (b) Forward UQ *(Source: E. Walker, J. Sun, and J. Chen).*

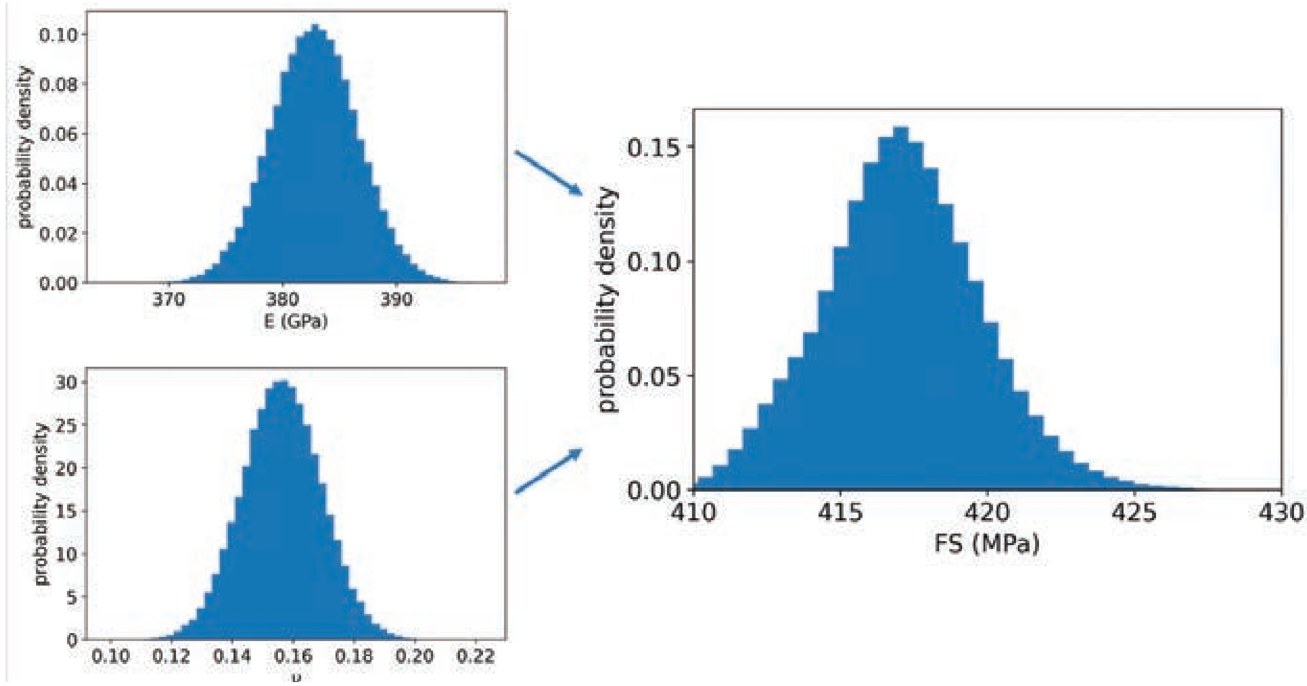


Figure 5. A Forward UQ for FS at 1400 °C *(Source: E. Walker, J. Sun, and J. Chen).*

Table 1. Uncertainty Shape Parameters of the Gaussian Distributions

DESCRIPTIVE STATISTIC	YOUNG'S MODULUS E (GPa)	POISSON'S RATIO *V*	FS (MPa)
Mean μ	382.8	0.156	417.03
Standard deviation σ	3.828	0.013	2.041

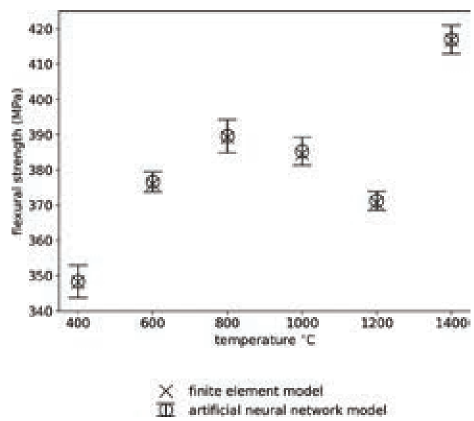


Figure 6. ANN and Physics Model Solved With FEM Prediction of FS vs. Temperature *(Source: E. Walker, J. Sun, and J. Chen).*

framework to predict the mechanical strength of α-SiC within seconds during each design iteration. The uncertainty has also been quantified at a standard deviation of 2.041 MPa. The entire training data set was used for the UQ conducted here. More training data will help stabilize the surrogate model.

The ANN is effective at capturing FS given its nonlinear dependence on temperature. Figure 6 displays the mean FS of the ANN as a function of temperature, labeled as circles, and its 95% confidence interval.

The thermomechanical fracture model solved via FEM, serving as the benchmark data, is overlaid with x's. The 95% confidence interval of the ANN is $\pm 2\sigma$ from the mean. The uncertainty of the ANN is temperature dependent—at its maximum, $\sigma = 2.36$ (MPa) at 800 °C; at its minimum, $\sigma = 1.36$ (MPa) at 1200 °C.

All distributions for the inputs in Table 1, namely Young's modulus and Poisson's ratio, are assumed as Gaussian and functions of temperature (equations 17 and 18). When defining the pdf of Young's modulus, three standard deviations are set to 25% of the mean. The same procedure is taken for defining the pdf of Poisson's ratio except for using 3% of the mean. To validate the smooth interpolation between training data temperatures, predictions are made and plotted in Figure 7.

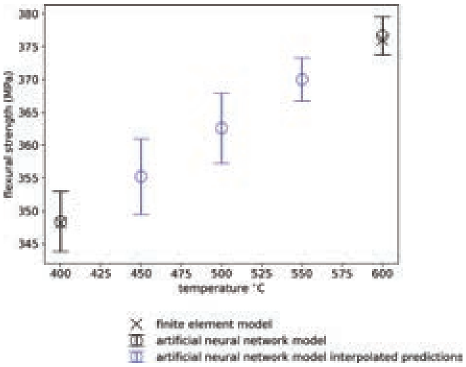


Figure 7. Interpolation Prediction of FS vs. Temperature *(Source: E. Walker, J. Sun, and J. Chen).*

The choice of three standard deviations in this work contrasts against absolute bounds because the Gaussian distribution used here has no absolute bounds. Munro [4] has the uncertainty at ±25% for Young's modulus and ±3% for Poisson's ratio as absolute uncertainty bounds. However, a Gaussian distribution has infinitely long bounds. It should be noted that 99.5% of the samples are within $\pm 3\sigma$ of the mean for a Gaussian pdf. Therefore, the choice of three standard deviations depends mostly on the literature precedent of uncertainties in α-SiC properties. The samples of the inputs are passed through the ANN, with FS samples as the output. The mean and standard deviation of the FS samples are taken at each temperature. The temperature reliance is from the oxidation of the surface [39] that increases FS.

CONCLUSIONS

The surrogate model shown here will prove useful for designing TPS's of hypersonic weapons. If manufacturing defects persist, there will be a need for UQ; this is enabled through the surrogate model. The surrogate model in this work is an ANN, and it has been cross-validated for accuracy. The results from the ANN cross-validation show that the training data set must contain enough of the nonlinear behavior of FS to exhibit reliable accuracy. The presented ANN is also used in an effective temperature-dependent UQ demonstration as a surrogate model. The reliability and accuracy of ANNs should remain in the parameter space of the training data. This study arrives at a similar conclusion. Rodríguez-Sánchez [40] used an ANN to model low-velocity impact force in a thermoplastic elastomer material (the ANN is accurate to 1% within the region of the training data).

Without UQ, designers of hypersonic weapons have no hint of the risk of their TPS's lacking FS during flight. The computation time of the presented work is greatly improved from the physics-based simulation. The parallelized physics-based simulation running on a cluster requires tens of minutes to complete a case. Run on a single processor, the ANN only requires seconds to train. An ANN prediction after training is faster—less than a second. Prediction of one million parameter samples requires seconds to run. With the modeling advancement shown here, the iterative step has been made in designing hypersonic weapons.

"Without UQ, designers of hypersonic weapons have no hint of the risk of their TPS's lacking FS during flight.

Questions regarding the future of ML in thermal protection systems modeling depend on how much physics is modeled by the ANN. For a separate application, As'ad et al. used physics-based constraints on an ANN to model stress and strain of elastic materials [41]. Those constraints include dynamic stability, objectivity, and consistency. ■

ACKNOWLEDGMENTS

Funding for this project was provided, in part, by LIFT, the Detroit-based national manufacturing innovation institute operated by the American Lightweight Materials Manufacturing Innovation Institute, a Michigan-based nonprofit, 501(c)3 as part of ONR Grant N00014-21-1-2660.

The authors are grateful to Dr. Amberlee Haselhuhn and Dr. David Hicks for the constructive discussions.

REFERENCES

[1] Sayler, K. M. "Hypersonic Weapons: Background and Issues for Congress." Congressional Research Service, 2023.

[2] Fleeman, E. "A Historical Overview of a Half Century of U.S. Missile Development." *DSIAC Journal*, vol. 3, no. 3, 2016.

[3] Lietha, M. N., and D. T. Motes III. "Understanding Thermal Protection Systems and the Development of the Hypersonic Speed Regime in the United States." Defense Systems Information Analysis Center (DSIAC) State-of-the-Art Report, DSIAC-2020-1326, December 2020.

[4] Munro, R. G. "Material Properties of a Sintered alpha-SiC." *Journal of Physical and Chemical Reference Data*, vol. 26, pp. 1195–1203, 1997.

[5] Alemany, M. M. E., F. Alarcón, F.-C. Lario, and J. J. Boj. "An Application to Support the Temporal and Spatial Distributed Decision-Making Process in Supply Chain Collaborative Planning." *Computers in Industry*, vol. 62, pp. 519–540, 2011.

[6] Souifi, A., Z. C. Boulanger, M. Zolghadri, M. Barkallah, and M. Haddar. "Uncertainty of Key Performance Indicators for Industry 4.0: A Methodology Based on the Theory of Belief Functions." *Computers in Industry*, vol. 140, 2022.

[7] Fan, H., W. T. Ang, and X. Wang. "On the Effective Property of a Micro-Cracked and a Microscopically Curved Interface Between Dissimilar Materials." *Forces in Mechanics*, vol. 7, 2022.

[8] Boljanović, S., and A. Carpinteri. "Modelling of the Fatigue Strength Degradation Due to a Semi-Elliptical Flaw." *Forces in Mechanics*, vol. 4, 2021.

[9] Dong, G., C. Lang, C. Li, and L. Zhang. "Formation Mechanism and Modelling of Exit Edge-Chipping During Ultrasonic Vibration Grinding of Deep-Small Holes of Microcrystalline-Mica Ceramics." *Ceramics International*, vol. 46, no. 8, part B, pp. 12458–12469, 2020.

[10] Knight, T. "Small Particle Effects on Hypersonic and Subsonic Flight Vehicles." DSIAC Technical Inquiry Response Report, DSIAC-BCO-2021-178, May 2021.

[11] Behnam, A., T. J. Truster, R. Tipireddy, M. C. Messner, and V. Gupta. "Uncertainty Quantification Framework for Predicting Material Response with Large Number of Parameters: Application to Creep Prediction in Ferritic-Martensitic Steels Using Combined Crystal Plasticity and Grain Boundary Models." *Integrating Materials and Manufacturing Innovation*, vol. 11, pp. 516–531, 2022.

[12] Saunders, R., A. Rawlings, A. Birnbaum, A. Iliopoulos, J. Michopoulos, D. Lagoudas, et al. "Additive Manufacturing Melt Pool Prediction and Classification via Multifidelity Gaussian Process Surrogates." *Integrating Materials and Manufacturing Innovation*, vol. 11, pp. 497–515, 2022.

[13] Nitta, K. H., and C. Y. Li. "A Stochastic Equation for Predicting Tensile Fractures in Ductile Polymer Solids." *Physica A: Statistical Mechanics and its Applications*, vol. 490, pp. 1076–1086, 2018.

[14] Štubňa, I., P. Šín, A. Trník, and L. Vozár. "Measuring the Flexural Strength of Ceramics at Elevated Temperatures – An Uncertainty Analysis." *Measurement Science Review*, vol. 14, pp. 35–40, 2014.

[15] Panchal, J. H., S. R. Kalidindi, and D. L. McDowell. "Key Computational Modeling Issues in Integrated Computational Materials Engineering." *Computer-Aided Design*, vol. 45, pp. 4–25, 2013.

[16] Shi, J., T. Cook, and J. Matlik. "Integrated Computational Materials Engineering for Ceramic Matrix Composite Development." The 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference, 2012.

[17] Sepe, R., A. Greco, A. D. Luca, F. Caputo, and F. Berto. "Influence of Thermo-Mechanical Material Properties on the Structural Response of a Welded Butt-Joint by FEM Simulation and Experimental Tests." *Forces in Mechanics*, vol. 4, 2021.

[18] Yaylaci, M. "Simulate of Edge and an Internal Crack Problem and Estimation of Stress Intensity Factor Through Finite Element Method." *Advanced Nano Research*, vol. 12, pp. 405–414, 2022.

[19] Becerra, A., A. Prabhu, M. S. Rongali, S. C. S. Velpur, B. Debusschere, and E. A. Walker. "How a Quantum Computer Could Quantify Uncertainty in Microkinetic Models." *The Journal of Physical Chemistry Letters*, vol. 12, no. 29, pp. 6955–6960, 2021.

[20] Becerra, A., O. H. Diaz-Ibarra, K. Kim, B. Debusschere, and E. A. Walker. "How a Quantum Computer Could Accurately Solve a Hydrogen-Air Combustion Model." *Digital Discovery*, vol. 1, pp. 511–518, 2022.

[21] Rumelhart, D. E., G. E. Hinton, and R. J. Williams. "Learning Representations by Back-Propagating Errors." *Nature*, vol. 323, pp. 533–536, 1986.

[22] Nair, V., and G. Hinton. "Rectified Linear Units Improve Restricted Boltzmann Machines." Proceedings of the 27th International Conference on Machine Learning, pp. 807–814, June 2010.

[23] Shanno, D. F. "Conditioning of Quasi-Newton Methods for Function Minimization." *Mathematics of Computation*, vol. 24, 1970.

[24] Kingma, D. P., and J. Ba. "Adam: A Method for Stochastic Optimization." International Conference on Learning Representations, 2015.

[25] Van Vinh, P., N. Van Chinh, and A. Tounsi. "Static Bending and Buckling Analysis of Bi-Directional Functionally Graded Porous Plates Using an Improved First-Order Shear Deformation Theory and FEM." *European Journal of Mechanics – A/Solids*, vol. 96, 2022.

[26] Cuong-Le, T., K. D. Nguyen, H. Le-Minh, P. Phan-Vu, P. Nguyen-Trong, and A. Tounsi. "Non-Linear Bending Analysis of Porous Sigmoid FGM Nanoplate via IGA and Nonlocal Strain Gradient Theory." *Advances in Nano Research*, vol. 12, pp. 441–455, 2022.

[27] Kuma, Y., A. Gupta, and A. Tounsi. "Size-Dependent Vibration Response of Porous Graded Nanostructure With FEM and Nonlocal Continuum Model." *Advances in Nano Research*, vol. 11, pp. 1–17, 2021.

[28] Alimirzaei, S., M. Mohammadimehr, and A. Tounsi. "Nonlinear Analysis of Viscoelastic Micro-Composite Beam With Geometrical Imperfection Using FEM: MSGT Electromagnetoelastic Bending, Buckling and Vibration Solutions." *Structural Engineering and Mechanics*, vol. 71, pp. 485–502, 2019.

[29] Miehe, C., F. Welschinger, and M. Hofacker. "Thermodynamically Consistent Phase-Field Models of Fracture: Variational Principles and Multi-Field FE Implementations." *International Journal for Numerical Methods in Engineering*, vol. 83, pp. 1273–1311, 2010.

[30] Miehe, C., L. Schänzel, and H. Ulmer. "Phase Field Modeling of Fracture in Multi-Physics Problems. Part I. Balance of Crack Surface and Failure Criteria for Brittle Crack Propagation in Thermo-Elastic Solids." *Computer Methods in Applied Mechanics and Engineering*, vol. 294, pp. 449–485, 2015.

[31] Zhang, X., C. Vignes, S. W. Sloan, and D. Sheng. "Numerical Evaluation of the Phase-Field Model for Brittle Fracture with Emphasis on the Length Scale." *Computational Mechanics*, vol. 59, pp. 737–752, 2017.

[32] Clayton, J. D., R. B. Leavy, and J. Knap. "Phase Field Modeling of Heterogeneous Microcrystalline Ceramics." *International Journal of Solids and Structures*, vol. 166, pp. 183–196, 2019.

[33] Sun, J., Y. Chen, J. J. Marziale, E. A. Walker, D. Salac, and J. Chen. "Damage Prediction of Sintered alpha-SiC Using Thermo-mechanical Coupled Fracture Model." *Journal of the American Ceramic Society*, accepted for publication, 2022.

[34] American Society for Testing and Materials (ASTM) C1161-18. "Standard Test Method for Flexural Strength of Advanced Ceramics at Ambient Temperature," 24 February 2023.

[35] Kendall, K. "Chapter 6 – Improving Fracture Mechanics (FM): Let's Get Back to Energy." In *Crack Control*, edited by K. Kendall, editor, Elsevier, pp. 137–163, 2021.

[36] Nagaraju, H. T., J. Nance, N. H. Kim, B. Sankar, and G. Subhash. "Uncertainty Quantification in Elastic Constants of SiCf/SiCm Tubular Composites Using Global Sensitivity Analysis." *Journal of Composite Materials*, vol. 57, pp. 63–78, 2023.

[37] Walker, E., J. Kammeraad, J. Goetz, M. T. Robo, A. Tewari, and P. M. Zimmerman. "Learning to Predict Reaction Conditions: Relationships Between

Solvent, Molecular Structure, and Catalyst." *Journal of Chemical Information and Modeling*, vol. 59, pp. 3645–3654, 2019.

[38] Walker, E. A., K. Ravisankar, and A. Savara. "CheKiPEUQ Intro 2: Harnessing Uncertainties From Data Sets, Bayesian Design of Experiments in Chemical Kinetics." *ChemCatChem*, vol. 12, no. 21, pp. 5401–5410, 2020.

[39] Arai, Y., R. Inoue, K. Goto, and Y. Kogo. "Carbon Fiber Reinforced Ultra-High Temperature Ceramic Matrix Composites: A Review." *Ceramics International*, vol. 45, pp. 14481–14489, 2019.

[40] Rodríguez-Sánchez, A. E. "An Artificial Neural Networks Approach to Predict Low-Velocity Impact Forces in an Elastomer Material." *Simulation*, vol. 96, pp. 551–563, 2020.

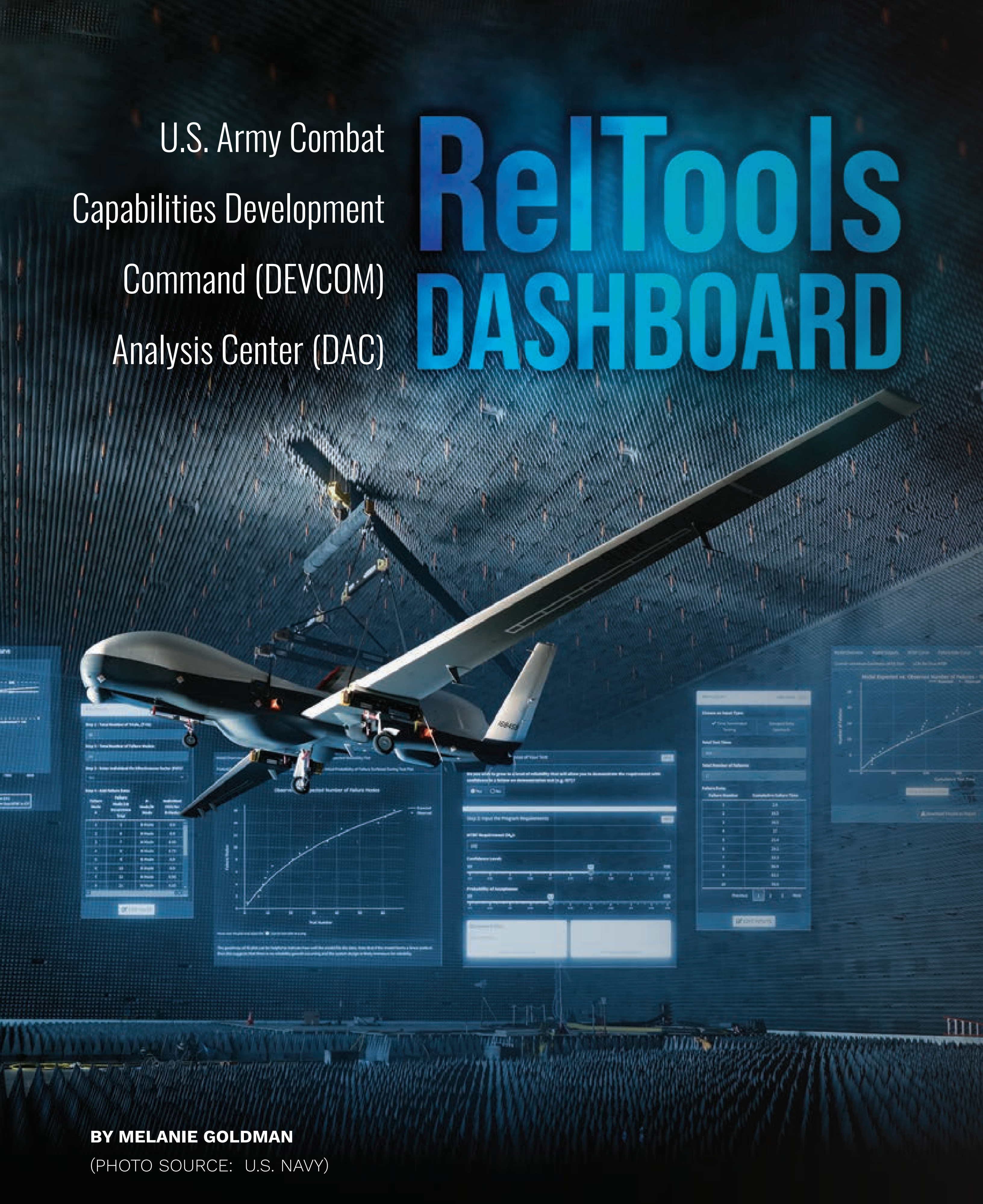
[41] As'ad, F., P. Avery, and C. Farhat. "A Mechanics-Informed Artificial Neural Network Approach in Data-Driven Constitutive Modeling." *International Journal for Numerical Methods in Engineering*, vol. 123, pp. 2738–2759, 2022.

BIOGRAPHIES

ERIC ALAN WALKER is a research scientist who applies artificial intelligence to new technologies and scientific challenges. His background is in chemical engineering, with a focus on artificial intelligence and data science. In recent years, he has learned integrated computational materials engineering and manufacturing from his colleagues. Dr. Walker holds a Ph.D. in chemical engineering.

JASON SUN is a third-year Ph.D. candidate majoring in aerospace engineering at the University at Buffalo (UB) under the advisement of Dr. James Chen. His research focuses on computational mechanics, multiscale modeling, and integrated computational materials engineering development for ceramic matrix composites. Mr. Sun holds a B.S. in mechanical and aerospace engineering and a B.A. in mathematics from UB.

JAMES M. CHEN is an associate professor in the Department of Mechanical and Aerospace Engineering at UB. He has published ~50 peer-reviewed journal articles in multiscale computational mechanics, theoretical and computational fluid dynamics, and atomistic simulation for thermoelectromechanical coupling. His research at the Multiscale Computational Physics Lab has been recognized by numerous media outlets, including a feature article in *Aerospace Testing International* (UK) and a radio show in Austria. He is a fellow of the American Society of Mechanical Engineers and an associate fellow of the American Institute of Aeronautics and Astronautics. Dr. Chen holds a B.S. in mechanical engineering from National Chung-Hsing University, an M.S. in applied mechanics from National Taiwan University, and a Ph.D. in mechanical and aerospace engineering from The George Washington University.



U.S. Army Combat Capabilities Development Command (DEVCOM) Analysis Center (DAC)

RelTools DASHBOARD

BY MELANIE GOLDMAN

(PHOTO SOURCE: U.S. NAVY)

BACKGROUND

Reliability is the probability that an item will perform its intended function for a specified period under stated conditions. It has a significant impact on operating and sustainment costs within the U.S. Department of Defense (DoD), which typically represent 70% of the program's total life-cycle costs [1]. System reliability also directly affects system availability and life-cycle costs since a system with poor reliability may require an increase in maintenance labor, repair material, and spares. Thus, improving reliability can result in significant savings in the support costs for any major weapon system.

Reliability growth is the improvement in reliability over time due to improvements in the product's design or the manufacturing process [2]. Reliability growth consists of two main areas—planning and assessment. Reliability growth assessment uses data from testing to estimate the improvement in reliability, and it can be further broken down into reliability growth tracking and projection. Reliability growth planning uses assessment model forms, along with various management metrics, to provide a plan for improving reliability prior to any testing. DAC developed reliability growth planning and assessment models, which can assist program managers in meeting the system reliability requirements and generate reliability growth curves for planning, tracking, and projection. In addition to providing reliability models, DAC also provides analytic support to defense programs and helps develop reliability-related policy, guidance, standards, methods, and training.

Reliability growth planning uses assessment model forms, along with various management metrics, to provide a plan for improving reliability prior to any testing.

RELTOOLS DASHBOARD

DAC, formally known as the U.S. Army Materiel Systems Analysis Activity (AMSAA), initially developed a collection of key reliability models and tools in Microsoft Excel and distributed them upon request to U.S. government personnel and their DoD contractors. The collection has been improved and modernized by coding the models in the programming language R and wrapping them in a Shiny app. The brand-new, highly interactive container for the models and tools is called the RelTools Dashboard (shown in Figure 1). The dashboard contains a wide range of models for reliability growth planning, tracking, and projection as well as tools for test planning and quick requirements calculations. It also hosts a downloadable Reliability Scorecard, which aids in assessing the reliability risks of a developmental acquisition program. As seen in Figure 1, the models are listed in the sidebar on the left-hand side and broken down by category. The dashboard can be accessed on the Web by a uniform resource locator [3] and is available free of charge to U.S. government

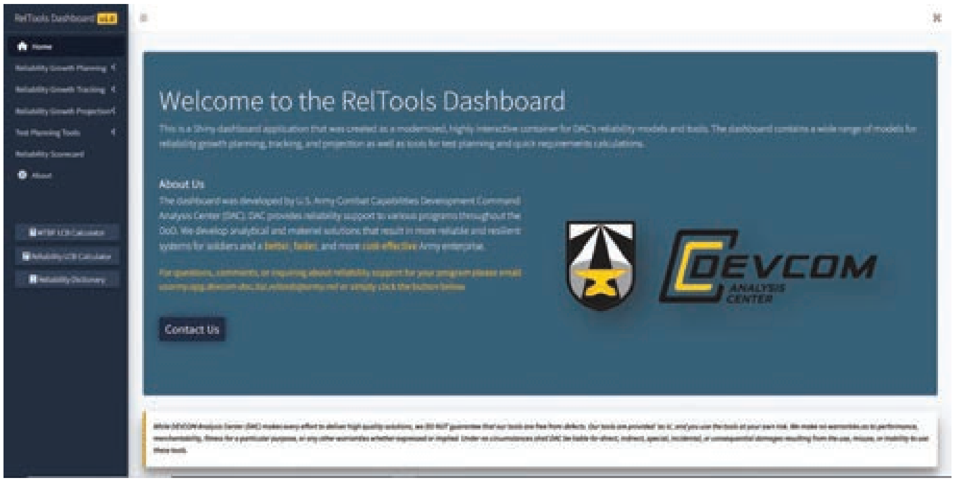


Figure 1. RelTools Dashboard Home Page *(Source: DAC).*

personnel and DoD contractors. The only requirements for access are a government network connection and a common access card.

RELIABILITY TEST PLANNING

When choosing the duration for reliability demonstration during an initial operational test (IOT), it is critical for program managers to consider the reliability requirement and the desired confidence level and probability of acceptance. To aid in this process, two reliability test planning tools are available: (1) the IOT Planning Tool – Continuous (IPT-C), which is for continuous-use systems, and (2) the IOT Planning Tool – Discrete (IPT-D), which is for discrete-use systems. A continuous-use system refers to systems where usage is measured on a continuous scale, such as hours, miles, etc. A discrete-use system refers to a system where usage is measured in terms of discrete trials, such as shots from guns, rockets, or missile systems. Both IPT-C and IPT-D allow program management to effectively plan for a fixed-length, fixed-configuration demonstration test, such as the IOT. These tools are located in the Test Planning Tools tab of the RelTools Dashboard. Screenshots from the IPT-C are shown in Figure 2.

The purpose of the IPT-C and IPT-D is to help the user choose an appropriate IOT profile to aid in developing a reliability program plan. Both IPT models contain two tools—a Test Length Planner and Operating Characteristic (OC) Curve Plotter. Their only difference is in the calculations used for the specific type of systems they cover. The

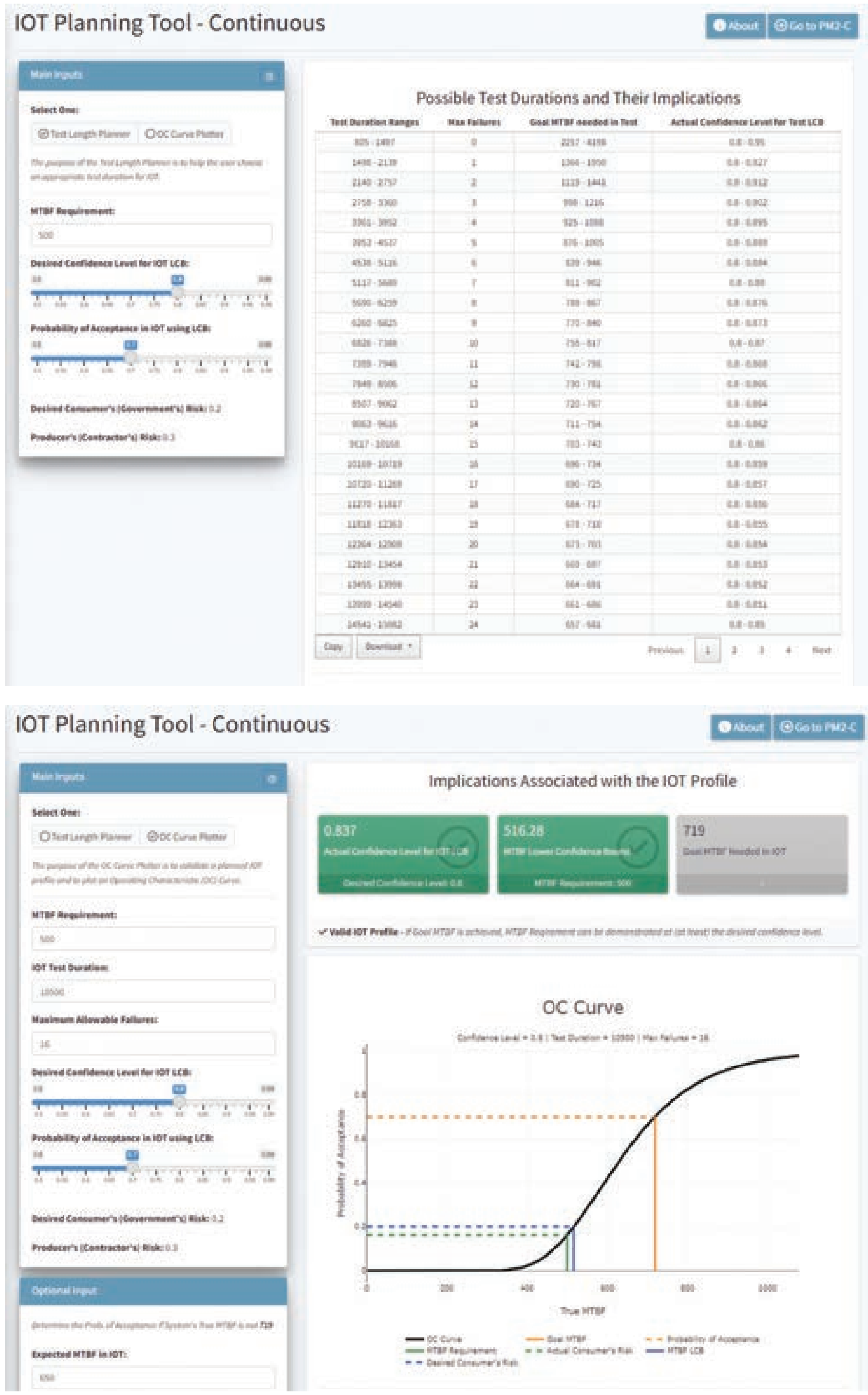


Figure 2. IPT-C Example *(Source: DAC).*

> ***The purpose of the IPT-C and IPT-D is to help the user choose an appropriate IOT profile to aid in developing a reliability program plan.***

Test Length Planner, shown on the left in Figure 2, displays possible IOT test durations based on the mean time between failure (MTBF) requirement, desired confidence level, and probability of acceptance. The OC Curve Plotter, shown on the right in Figure 2, is used to validate a planned IOT profile and plot an OC curve. The OC curve shows how the true underlying reliability of the system impacts the probability of successfully passing the demonstration test. Together, these tools can aid in conducting trade-off analyses involving the reliability demonstration test lengths, reliability requirements, and the associated statistical risks when planning reliability demonstration tests.

RELIABILITY GROWTH PLANNING

Reliability growth occurs by identifying failure modes and developing corrective actions to mitigate the failure modes. Reliability Growth Planning Curves (RGPCs) assist program managers in managing this process by determining program schedules, allocating resources, and defining the realism of the test program in achieving the required reliability. The planning curve is constructed early in the program's life cycle when little-to-no reliability data is available, and it provides an indication of the reliability that can be expected during different stages of the program's development. The planned reliability values in the planning curve serve as a basis for evaluating future reliability growth progress during reliability growth testing. An example of a RGPC is shown in Figure 3. The idealized curve, shown in black, portrays the expected overall reliability growth pattern across test phases. Since failure modes are not found and corrected instantaneously during testing, the RGPC uses a series of MTBF targets to represent the actual MTBF goals for the system during each test phase throughout the test program [2]. These are represented by the sequence of colored steps in the plot. The end of the planning curve depicts the reliability target necessary to successfully demonstrate the reliability requirement in a reliability demonstration test. This target can be found using the IPT-C previously discussed.

Reliability growth planning should consider the initial and goal reliability targets, test phases, corrective action periods, and reliability thresholds (interim goals to be achieved following corrective action periods). Reliability

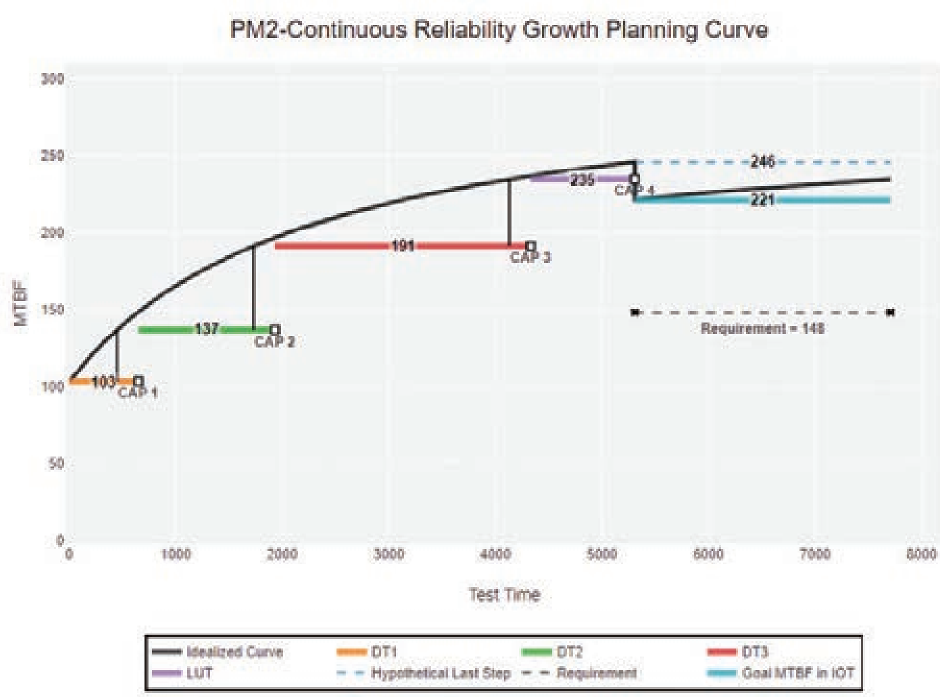


Figure 3. Example of Reliability Growth Planning Curve *(Source: DAC).*

growth planning should also include realistic management metrics, such as management strategy (MS) and fix effectiveness factors (FEFs). MS is the assumed proportion of the initial failure rate that will be addressed via corrective action during the test-fix-test process. FEF is the fractional reduction in the failure rate for a failure mode after a corrective action is applied [2].

The RGPC should be developed using reliability growth planning models, such as the Planning Model Based on Projection Methodology (PM2)-Continuous (PM2-C) or PM2-Discrete (PM2-D). The purpose of PM2-C and PM2-D is to aid in constructing a reliability growth planning curve (similar to Figure 3) over a developmental test (DT) program useful to program management. It serves as a baseline against which reliability assessments can be compared and can highlight the need to management when reallocation of resources is necessary. Both models are found under the Reliability Growth Planning tab of the RelTools Dashboard, and screenshots of the main inputs of these models are shown in Figures 4 and 5. The main inputs are broken down in a series of steps and include the following:

- MTBF or reliability requirement
- Confidence level
- Probability of acceptance
- IOT test duration or number of trials
- Assumed DT to IOT degradation factor
- Initial MTBF or reliability
- MS
- Average FEF
- DT schedule

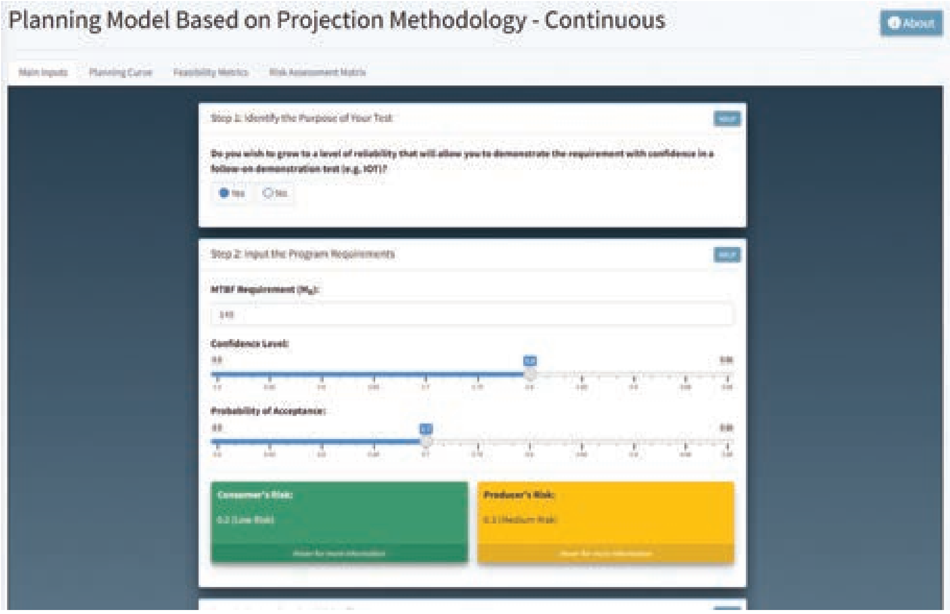


Figure 4. PM2-C Main Inputs *(Source: DEVCOM Analysis Center).*

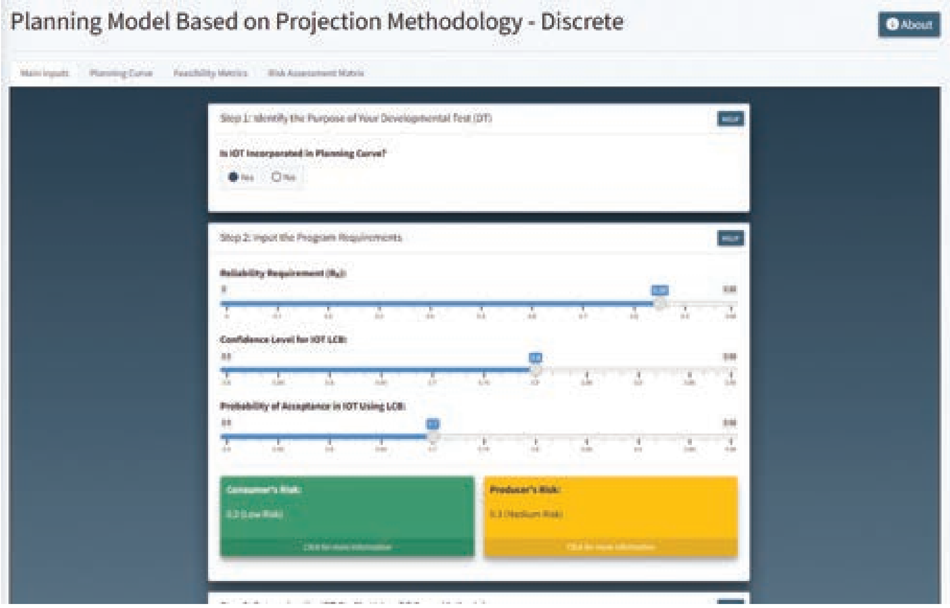


Figure 5. PM2-D Main Inputs *(Source: DAC).*

RELIABILITY GROWTH TRACKING

Reliability Growth Tracking models estimate the reliability of a system in a DT program by evaluating the reliability growth that results from incorporating design and quality fixes into the system during a test event. The purpose of reliability growth tracking is to determine the amount of reliability growth occurring during the test and estimate the demonstrated reliability at the end of the test phase. This assists management in determining if the program is progressing as planned, better than planned, or not as well as planned. When a program is not progressing as planned, a new reliability strategy may need to be developed. Reliability growth tracking provides a statistical estimate of the system reliability at a given time based on observed test data. Tracking models should be used to update a system's reliability based on actual test data.

The Reliability Growth Tracking Model for Continuous data (RGTMC) is a tool for assessing the improvement (growth) in the system's reliability, within a single test phase during development, for which usage is continuously measured. This model is useful for monitoring a system's reliability in a DT program by evaluating the reliability growth resulting from incorporating design and quality fixes into the system during the test program. The model may utilize data when exact failure times are known, along with data when failure times are only known to an interval (grouped data).

RGTMC is found under the Reliability Growth Tracking tab of the RelTools Dashboard, and a screenshot of the model is shown in Figure 6. The figure depicts the model after main inputs have been entered and calculations have been performed. The left-hand side shows the inputs that were entered into the model, including total test time, total number of failures, and failure occurrence times. The right-hand side shows an example of the expected vs. observed number of failures plot generated by RGTMC, which can be used to determine how well the model fits the data. Additional model output, such as the estimated MTBF at the end of the test, is available in the tabs on the top right of the screen.

RELIABILITY GROWTH PROJECTION

Reliability growth projection is an additional approach for determining if a program is on track to meeting its reliability requirements. Reliability Growth Projection models provide an estimate of the reliability at a current or future milestone based on planned and/or implemented fixes, assessed fix effectiveness, and a statistical estimate of the problem mode rates of occurrence. Projections determine the system's reliability before and after corrective actions have been implemented, including future delayed, corrective actions. These results can be used to help program managers

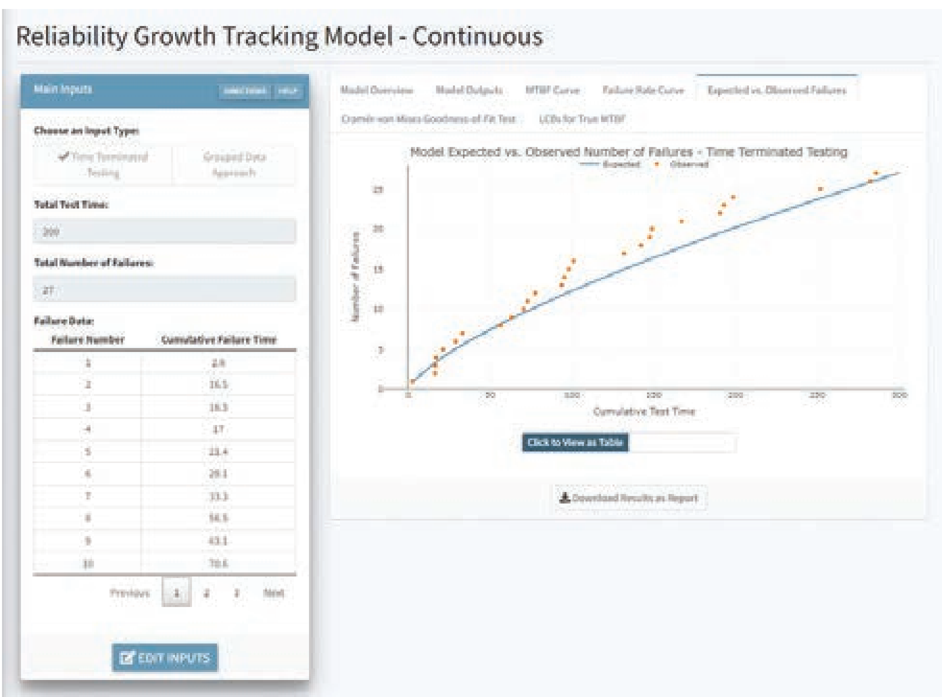


Figure 6. RGTM-C Example *(Source: DAC)*.

> *Projections determine the system's reliability before and after corrective actions have been implemented, including future delayed, corrective actions.*

decide how to allocate resources prior to entering the next test phase. When a program is experiencing reliability problems, projections can be used to investigate alternative test plans. This can be done using either the AMSAA Maturity Projection Model (AMPM) [4] or the AMSAA Discrete Projection Model (ADPM) [5], which can be found in the Reliability Growth Projection tab of the RelTools Dashboard.

The purpose of the AMPM is to provide an estimate of the projected reliability following the implementation of both delayed and nondelayed fixes for *continuous systems*. The model also provides estimates of the following important reliability growth metrics:

- Initial failure intensity
- Growth potential failure intensity
- Expected number of failure modes surfaced
- Percent of the initial failure intensity surfaced
- Rate of occurrence of new failure modes

The purpose of the ADPM is to provide an estimate of the projected reliability following the implementation of both delayed and nondelayed fixes for *discrete systems*. The model also provides estimates of the following important reliability growth metrics:

- Reliability growth potential
- Expected number of failure modes surfaced
- Probability of a new failure mode occurring
- Fraction surface of the system's initial probability of failure

Screenshots from AMPM and ADPM are shown in Figures 7 and 8, respectively. These figures depict both models after main inputs have been entered and calculations have been performed. In Figure 7, the left-hand side shows the inputs that were entered into the model, and the right-hand side shows the observed failure modes vs. the prior predicted, cumulative number of failure modes plot. In Figure 8, the left-hand side shows the inputs that were entered into the model, and the right-hand side shows the observed vs. expected number of failure modes plot.

RELIABILITY SCORECARD

In addition to reliability growth models and test planning tools, the RelTools Dashboard also contains the DAC Reliability Scorecard. This scorecard replaces the AMSAA Reliability Scorecard [6] and AMSAA Software Reliability Scorecard [7]. The purpose of the DAC Reliability Scorecard is to evaluate a program's

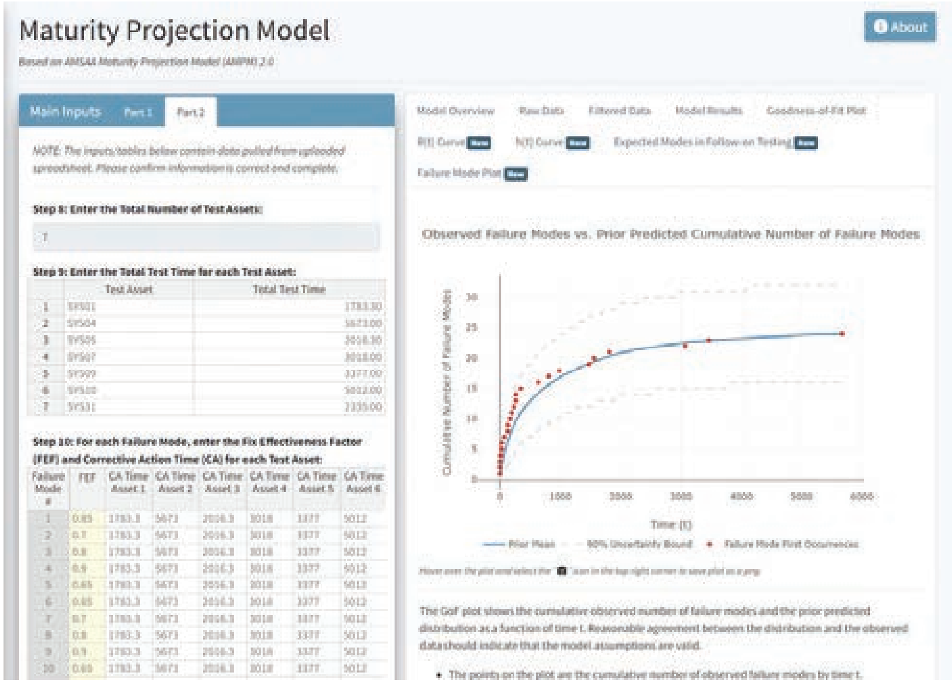


Figure 7. AMPM Example *(Source: DAC)*.

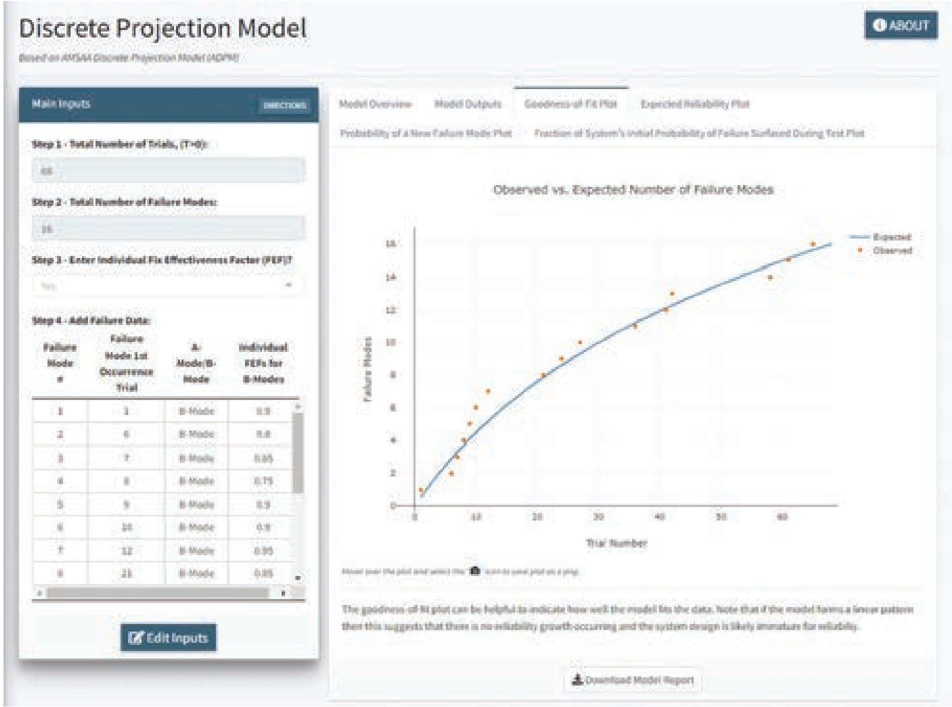


Figure 8. ADPM Example *(Source: DAC)*.

> ***The purpose of the DAC Reliability Scorecard is to evaluate a program's planned and completed reliability activities.***

planned and completed reliability activities. This is to ensure reliability best practices are being implemented while also identifying areas that may need improvement. It can be applied to a program at any phase in the acquisition life cycle and to systems that are hardware-intensive, software-intensive, or a combination of both. The Scorecard can be found in the Reliability Scorecard tab and is available as a downloadable Excel file to be used separately from the Dashboard itself. The spreadsheet consists of the following eight categories:

1. Program Plan and Schedule
2. Develop and Design Team
3. Requirements and Goals
4. Design Process and Considerations
5. Modeling and Analysis
6. Testing
7. Supply Chain
8. Fielding and Sustainment

These eight categories contain one or more elements, resulting in 32 total elements. Based on the rating criteria associated with each element, the user will assign each element one of the following ratings from a drop-down list: Not Achieved, Partially Achieved/ Needs Improvement, Fully Achieved, or Not Applicable. Once all elements are rated, the scorecard generates a summary of the scorecard ratings, which can assess the adequacy of a reliability program and highlight the areas that require improvement.

CONCLUSIONS

System reliability is a key component of maximizing system availability and minimizing life-cycle costs. Overall, an improvement in reliability can result in significant savings in the support costs for any major weapon system. The models and tools available in the RelTools Dashboard can be used to improve reliability by assisting management in meeting system reliability requirements. To access the RelTools Dashboard, go to https://apps.dse.futures.army.mil/RelToolsDashboard/. For any questions about the dashboard or any of the models, please contact usarmy.apg.devcom-dac.list.reltools@army.mil. ■

REFERENCES

[1] U.S. Government Accountability Office (GAO). "Weapon System Sustainment: Selected Air Force and Navy Aircraft Generally Have Not Met Availability Goals, and DoD and Navy Guidance Need to Be Clarified." GAO-18-678, Washington, D.C., 10 September 2018.

[2] U.S. Army Materiel Systems Analysis Activity. "MILHDBK-189C: Reliability Growth Management." Aberdeen Proving Ground, MD, 14 June 2011.

[3] DEVCOM Analysis Center. "RelTools Dashboard." https://apps.dse.futures.army.mil/RelToolsDashboard, 26 October 2022.

[4] Wayne, M., and M. Modarres. "A Bayesian Model for Complex System Reliability Growth Under Arbitrary Corrective Actions." *IEEE Transactions on Reliability*, vol. 64, no. 1, pp. 206–220, doi: 10.1109/TR.2014.2337072, March 2015.

[5] Hall, J. B., P. M. Ellner, and A. Mosleh. "Reliability Growth Management Metrics and Statistical Methods for Discrete-Use Systems." *Technometrics*, vol. 52, no. 4, pp. 379–389, doi: 10.1198/tech.2010.08068, 2010.

[6] Shepler, M. H., and N. Welliver. "New Army and DoD Reliability Scorecard." Proceedings – Annual Reliability and Maintainability Symposium (RAMS), San Jose, CA, pp. 1–5, doi: 10.1109/RAMS.2010.5448062, 2010.

[7] Bernreuther, D., and T. Pohland. "Improved Software Reliability by Application of Scorecard Reviews." The 2013 Proceedings Annual Reliability and Maintainability Symposium (RAMS), Orlando, FL, pp. 1–5, doi: 10.1109/RAMS.2013.6517767, 2013.

BIOGRAPHY

MELANIE GOLDMAN is a member of the Center for Reliability Growth Team at DAC. She holds a bachelor's degree in mathematics from the University of Delaware and is currently pursuing her master's degree in applied and computational mathematics from Johns Hopkins University.

DSIAC JOURNAL

The Defense Systems Information Analysis Center (DSIAC) is a component of the U.S. Department of Defense's (DoD's) Information Analysis Center (IAC) enterprise, serving the defense enterprise of DoD and federal government users and their supporting academia and industry partners.

WWW.DSIAC.ORG

CONNECT WITH US ON SOCIAL MEDIA