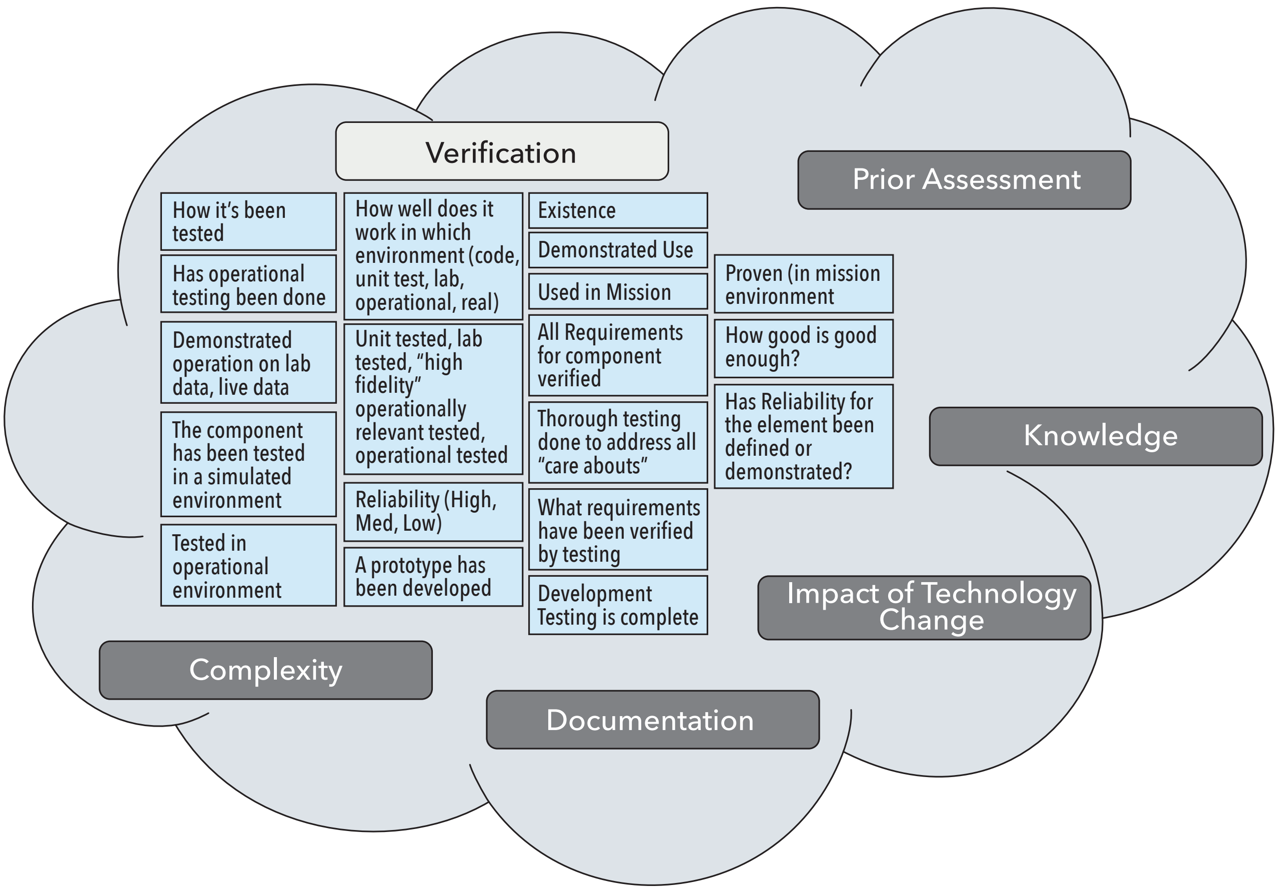
INSIGHT

This Issue's Feature:

Uncertainty in the Engineering of Systems



A snapshot of the Bayesian network categories and specifically the sub-categories for verification

Illustration credit: from the article
Applying Bayesian Networks to TRL Assessments – Innovation in Systems Engineering
by Marc F. Austin, Virginia Ahalt, Erin Doolittle, Cheyne Homberger, George A. Polacek, and Donald M. York (page 47)

DECEMBER 2024
VOLUME 27 / ISSUE 6

INCOSE

A PUBLICATION OF THE INTERNATIONAL COUNCIL ON SYSTEMS ENGINEERING

MISSOURI S&T

Earn your M.S. in Systems Engineering online today!

Modular approach

No thesis required

30 credit hours

There's a reason Missouri University of Science and Technology leads the way in online systems engineering graduate education: We've been practicing and perfecting the art of using technology to engage and inspire for a long time. With faculty who teach based on real-world systems engineering experience, a global network of successful alumni and a curriculum that sets standards across the industry, Missouri S&T is the place for you.

online.mst.edu

INSIGHT

A PUBLICATION OF THE INTERNATIONAL COUNCIL ON SYSTEMS ENGINEERING

DECEMBER 2024 VOLUME 27/ ISSUE 6

Inside this issue

About This Publication

INFORMATION ABOUT INCOSE

INCOSE's membership extends to over 25,000 members and CAB associates and more than 200 corporations, government entities, and academic institutions. Its mission is to share, promote, and advance the best of systems engineering from across the globe for the benefit of humanity and the planet. INCOSE charters chapters worldwide, includes a corporate advisory board, and is led by elected officers and directors.

For more information, click here:
The International Council on Systems Engineering
(www.incose.org)

OVERVIEW

INSIGHT is the magazine of the International Council on Systems Engineering. It is published six times per year and features informative articles dedicated to advancing the state of practice in systems engineering and to close the gap with the state of the art. ***INSIGHT*** delivers practical information on current hot topics, implementations, and best practices, written in applications-driven style. There is an emphasis on practical applications, tutorials, guides, and case studies that result in successful outcomes. Explicitly identified opinion pieces, book reviews, and technology roadmapping complement articles to stimulate advancing the state of practice. ***INSIGHT*** is dedicated to advancing the INCOSE objectives of impactful products and accelerating the transformation of systems engineering to a model-based discipline.
Topics to be covered include resilient systems, model-based systems engineering, commercial-driven transformational systems engineering, natural systems, agile security, systems of systems, and cyber-physical systems across disciplines and domains of interest to the constituent groups in the systems engineering community: industry, government, and academia. Advances in practice often come from lateral connections of information dissemination across disciplines and domains. ***INSIGHT*** will track advances in the state of the art with follow-up, practically written articles to more rapidly disseminate knowledge to stimulate practice throughout the community.

Editor-In-Chief William Miller
insight@incose.net +1 908-759-7110
Layout and Design Chuck Eng
chuck.eng@comcast.net
Member Services INCOSE Administrative Office
info@incose.net +1 858 541-1725

ARTICLE SUBMISSION insight@incose.net

Publication Schedule. ***INSIGHT*** is published six times per year. Issue and article submission deadlines are as follows:

- February 2025 issue – 1 November 2024
- April 2025 issue – 2 January 2025
- June 2025 issue – 1 March 2025
- August 2025 issue – 1 May 2025
- October 2025 – 1 July 2025
- December 2025 – 1 September 2025

For further information on submissions and issue themes, visit the INCOSE website: www.incose.org

 ISSN 2156-485X; (print) ISSN 2156-4868 (online)

ADVERTISE

Readership
INSIGHT reaches over 25,000 members and CAB associates and uncounted employees and students of more than 130 CAB organizations worldwide. Readership includes engineers, manufacturers/purchasers, scientists, research and development professionals, presidents and chief executive officers, students, and other professionals in systems engineering.

Issuance	Circulation
2024, Vol 27, 6 Issues	100% Paid

Contact us for Advertising and Corporate Sales Services
We have a complete range of advertising and publishing solutions professionally managed within our global team. From traditional print-based solutions to cutting-edge online technology the Wiley-Blackwell corporate sales service is your connection to minds that matter. For an overview of all our services please browse our site which is located under the Resources section. Contact our corporate sales team today to discuss the range of services available:

- Print advertising for non-US journals
- Email Table of Contents Sponsorship
- Reprints
- Supplement and sponsorship opportunities
- Books
- Custom Projects
- Online advertising

Click on the option below to email your enquiry to your nearest office:

- Asia and Australia corporatesalesaustralia@wiley.com
- Europe, Middle East and Africa (EMEA) corporatesaleseurope@wiley.com
- Japan corporatesalesjapan@wiley.com
- Korea corporatesaleskorea@wiley.com

USA (also Canada, and South/Central America):

- Healthcare Advertising corporatesalesusa@wiley.com
- Science Advertising Ads_sciences@wiley.com
- Reprints Commercialreprints@wiley.com
- Supplements, Sponsorship, Books and Custom Projects busdev@wiley.com

Or please contact: Marcom@incose.net

CONTACT

Questions or comments concerning:

Submissions, Editorial Policy, or Publication Management
Please contact: William Miller, Editor-in-Chief
insight@incose.net

Advertising—*please contact:*
Marcom@incose.net

Member Services – ***please contact:*** info@incose.org

ADVERTISER INDEX December Volume 27-6

CORPORATE ADVISORY BOARD — MEMBER COMPANIES

FROM THE EDITOR-IN-CHIEF

William Miller, insight@incose.net

We are pleased to publish the December 2024 *INSIGHT* published cooperatively with John Wiley & Sons as the systems engineering practitioners' magazine. The *INSIGHT* mission is to provide informative articles on advancing the practice of systems engineering as the state-of-the-art advances as evidenced in *Systems Engineering*, the Journal of INCOSE also published by Wiley, as well as papers presented at symposia and conferences by INCOSE and in the broader systems community.

The focus of this December issue of *INSIGHT* themed on uncertainty and Bayesian methods continues the systems engineering theoretical foundations and its impacts on practice in the April, August, and October 2024 issues of *INSIGHT* featuring the contributions of the "Bridge Team" (April) and the MBSE Patterns Working Group (August and October).

The imperative for systems engineering to address uncertainty is clearly called out in The *Systems Engineering Vision 2035: Engineering Solutions for a Better World* ©2021 by INCOSE (www.incose.org/publications/se-vision-2035).

The *Vision 2035* section Foundations (page 22):

> *Practicing systems engineers use a variety of analytical tools that are based on math and science. This requires competencies in the foundational math and science that is needed to analyze the systems of interest, and the enabling systems used to manufacture and support the systems.* ***The systems engineer also must understand how to use probability and statistics to understand risk and uncertainty****, and understand principles such as coupling and cohesion to manage systems complexity. For cyber-physical systems, a systems engineer of today must also have a basic understanding of control theory and communications.*

The *Vision 2035* section Specific Recommendations (to Systems Engineering Community) (pages 59-60):

Addressing Dynamic Change and ***Uncertainty***

- *Data standards are developed and adopted enabling effective data interconnection and exchange.*
- *Methods and tools for dealing with product variation and variability are widely adopted.*
- *Knowledge Management and incremental learning are integrated with systems engineering practices.*
- *Systems engineering incorporates dynamic feedback into solutions across the life cycle (such as Agile practices).*
- *Increasing technology assistance for human tasking is incorporated including automated workflows.*

Foundations and Research

- *New principles, phenomena, concepts, heuristics, and technologies are integrated with existing knowledge.*
- *Research to define and validate the systems engineering Theoretical Foundations is launched.*
- *Research on systems engineering practices, tools, and applications that address dynamic change and* ***uncertainty*** *is facilitated.*
- *Industry, government, and associations team with academia to further systems engineering research and incorporate systems engineering foundations into the curriculum.*
- *Systems engineering research encourages cross-disciplinary engagement to move towards integrated approaches.*

The imperative to address 'uncertainty' is a priority of the Future of Systems Engineering (FuSE) initiative to realize the *System Engineering Vision 2035*. Your editor also serves as the lead for the FuSE initiative and has been appointed the Assistant Director, FuSE in INCOSE Technical Operations to strengthen the relationship within our technical community. FuSE is empowered by the INCOSE Strategic Plan v1.0 (17 June 2024) Objective O.1 Advance systems engineering as the world's trusted authority and Key Result KR1.1 Satisfaction of/progress against future of systems engineering roadmap.

Your editor searched for 'uncertainty' in the ensemble of 4955 artifacts in the INCOSE content library of symposia papers, regional conferences, and webinars, retrieving 80 hits of which the great majority are symposia presentations and papers, and a few webinars. Most hits invoked the word 'uncertainty' without exposition; only a few went into the details of theory, methodology, practice, or examples. Searches for 'probability' yielded 35 hits and 'Bayesian' yielded 11 hits, the great majority of which are also without exposition. A stereotype of systems engineers is that we are uncomfortable with uncertainty from many anecdotal conversations with systems engineering colleagues in industry, government, and academia.

Yet, there is documented evidence that systems engineers have accounted for uncertainty in the engineering of systems beginning back in the early 20th Century

(CE). Goode and Machol expounded on accounting for probability (uncertainty) including the application of Bayes law, in *System Engineering: An Introduction to the Design of Large-Scale Systems* ©1957. An illustrative case study in their textbook is the telephone system, that is, the Bell System engineered by the Bell Telephone Laboratories, that achieved mandated quality of service objectives given dynamic uncertainties in demands for service, wide variations in environmental conditions in operations, dynamic regulatory constraints, technology constraints, and exponential growth that drove research and innovation in technologies progressing from manually operated switchboards to automatic systems progressing from electro-mechanical relays, to vacuum tubes, to discrete transistors, to integrated circuits, and to computer and software based digitization across the enterprise. Similarly, communications transmission progressed from copper wires, to multiplexed copper wires, to coax cable, to microwave radio, to fiber optics, to cellular radio frequency for mobility. All while maintaining interoperability across the ensemble of technologies! Hall's *A Methodology for Systems Engineering* ©1962 documented theories, methods, and applications innovated and practiced at Bell Labs for the Bell System, also accounting for uncertainty. Lastly, *Systems Architecture: Strategy and Product Development for Complex Systems* by Crawley, Cameron, and Selva ©2016 describes NASA's Apollo tradespace that assessed the probability of mission success compared to total mass for different mission concepts: direct (to the Moon and back), Earth-orbit rendezvous (EOR), lunar-orbit rendezvous (LOR), and EOR+LOR, with LOR selected as the concept (pages 311-317). For further exposition on risk and uncertainty by NASA, their *Probabilistic Risk Assessment Procedures Guide for NASA Managers and Practitioners* (NASA/SP-2011-3421, Second Edition, December 2011) is available online (https://ntrs.nasa.gov/citations/20120001369). The Artemis Missions Probabilistic Risk Assessment (PRA) & Reliability Assessment Overview (July 26, 2023) back to the Moon is also available online (https://ntrs.nasa.gov/api/citations/20230010735/downloads/Artemis%20PRA%20and%20Reliability%20Overview.pdf).

We lead the December *INSIGHT* with "Uncertainty Quantification (UQ) in Complex System of Systems (SoS) Modeling and Simulation (M&S) Environments" by Joseph Marvin, Thomas Whalen, Brad Morantz, Ray Deiotte, and Robert K. Garrett, Jr. The authors apply systems thinking to modeling and simulation (M&S) techniques to provide meaningful quantitative results in M&S of complex system of systems (SoSs) in the face of an environment filled with complex interacting uncertainties. They present a five-step statistical and parametric algorithm tool that addresses uncertainty quantification (UQ), proposing a quantitative approach to managing complex uncertainties across modeled interfaces using graph theory.

"Measuring the Uncertainty Impacts During the Systems Engineering Lifecycle" by David Flanigan and Jeffery Dixon explore a methodology to quantify uncertainty and the interdependencies of the uncertainty factors during development, including both internal and external factors and their contribution to the overall system uncertainty. An illustrative example is provided to illuminate the methodology.

"The ValXplore Method: Exploring Desirability, Feasibility and Viability of Business and System Design under Uncertainty" by Sonia Ben Hamida, Marija Jankovic, Alain Huet, and Jean-Claude Bocquet support decision-making in business and system design. ValXplore uses visual analysis and data analytics to perform rapid sensitivity and uncertainty analysis of a large number of design alternatives to explore uncertainties. They tested and validated the method on an industrial case study to assess the benefits and limits of the semi-reusability of a launch vehicle.

"Informing the Delineation of Input Uncertainty Space in Exploratory Modelling using a Heuristic Approach" by Enayat A. Moallemi, Sondoss Elsawah, and Michael J. Ryan propose a heuristic approach which informs the delineation of input uncertainties by screening the relevant model behavior in the solution space. An aircraft fleet management system is used to demonstrate the implementation of the approach in practice. The authors conclude that the delineation of input uncertainty space can be a way to control simulations in exploratory modelling and to enhance the efficiency of the exploration process and the confidence of the final results.

"Assessing the Impacts of Uncertainty Propagation to System Requirements by Evaluating Requirement Connectivity" by Alejandro Salado and Roshanak Nilchiani describe a requirement connectivity metric to evaluate the potential consequences that changing a requirement may have on a system with respect to fulfillment of other requirements. The metric is used to evaluate different cases in which requirements are changed due to triggering of uncertain events during a project life cycle.

"Applying Bayesian Networks to TRL Assessments – Innovation in Systems Engineering" by Marc F. Austin, Virginia Ahalt, Erin Doolittle, Cheyne Homberger, George A. Polacek, and Donald M. York argues for the use of a Bayesian network model to provide a mathematical method to consistently combine and validate the judgment of subject matter experts (SMEs) to increase the confidence in the determination of the readiness of system components and their technologies using technology readiness level (TRL) assessments to determine the maturity of technology readiness assessments (TRAs) and critical technology elements (CTEs) of a system as it moves forward in the system development life cycle.

"A Bayesian Approach for Estimating Complex System Reliability" by Ozge Doguc and Jose Emmanuel Ramirez-Marquez describes a holistic method using historical data about a system to be modeled as a Bayesian network (BN) that provides efficient techniques for automated construction of the BN model and estimation of the system reliability. Limited human intervention is sufficient for the process of BN construction and reliability estimation. This is in contrast to the time-consuming practice of having a group of specialists who are BN and domain experts build a system-specific BN that may lead to incorrect deductions for the specific system.

We hope you find *INSIGHT*, the practitioners' magazine for systems engineers, informative and relevant. Feedback from readers is critical to *INSIGHT*'s quality. We encourage letters to the editor at insight@incose.net. Please include "letter to the editor" in the subject line. *INSIGHT* also continues to solicit special features, standalone articles, book reviews, and op-eds. Please contact us at FuSE@incose.net if you are interested in contributing to our body of knowledge accounting for uncertainty in the engineering of systems. For information about *INSIGHT*, including upcoming issues, see https://www.incose.org/products-and-publications/periodicals#INSIGHT. For information about sponsoring *INSIGHT*, please contact the INCOSE marketing and communications director at marcom@incose.net. ■

Systems Engineering: The Journal of The International Council on Systems Engineering

Call for Papers

The *Systems Engineering* journal is intended to be a primary source of multidisciplinary information for the systems engineering and management of products and services, and processes of all types. Systems engineering activities involve the technologies and system management approaches needed for

- definition of systems, including identification of user requirements and technological specifications;
- development of systems, including conceptual architectures, tradeoff of design concepts, configuration management during system development, integration of new systems with legacy systems, integrated product and process development; and
- deployment of systems, including operational test and evaluation, maintenance over an extended life-cycle, and re-engineering.

Systems Engineering is the archival journal of, and exists to serve the following objectives of, the International Council on Systems Engineering (INCOSE):

- To provide a focal point for dissemination of systems engineering knowledge
- To promote collaboration in systems engineering education and research
- To encourage and assure establishment of professional standards for integrity in the practice of systems engineering
- To improve the professional status of all those engaged in the practice of systems engineering
- To encourage governmental and industrial support for research and educational programs that will improve the systems engineering process and its practice

The journal supports these goals by providing a continuing, respected publication of peer-reviewed results from research and development in the area of systems engineering. Systems engineering is defined broadly in this context as an interdisciplinary approach and means to enable the realization of successful systems that are of high quality, cost-effective, and trustworthy in meeting customer requirements.

The *Systems Engineering* journal is dedicated to all aspects of the engineering of systems: technical, management, economic, and social. It focuses on the life-cycle processes needed to create trustworthy and high-quality systems. It will also emphasize the systems management efforts needed to define, develop, and deploy trustworthy and high quality processes for the production of systems. Within this, *Systems Engineering* is especially concerned with evaluation of the efficiency and effectiveness of systems management, technical direction, and integration of systems. *Systems Engineering* is also very concerned with the engineering of systems that support sustainable development. Modern systems, including both products and services, are often very knowledge-intensive, and are found in both the public and private sectors. The journal emphasizes strategic and program management of these, and the information and knowledge base for knowledge principles, knowledge practices, and knowledge perspectives for the engineering of systems. Definitive case studies involving systems engineering practice are especially welcome.

The journal is a primary source of information for the systems engineering of products and services that are generally large in scale, scope, and complexity. *Systems Engineering* will be especially concerned with process- or product-line–related efforts needed to produce products that are trustworthy and of high quality, and that are cost effective in meeting user needs. A major component of this is system cost and operational effectiveness determination, and the development of processes that ensure that products are cost effective. This requires the integration of a number of engineering disciplines necessary for the definition, development, and deployment of complex systems. It also requires attention to the lifecycle process used to produce systems, and the integration of systems, including legacy systems, at various architectural levels. In addition, appropriate systems management of information and knowledge across technologies, organizations, and environments is also needed to insure a sustainable world.

The journal will accept and review submissions in English from any author, in any global locality, whether or not the author is an INCOSE member. A body of international peers will review all submissions, and the reviewers will suggest potential revisions to the author, with the intent to achieve published papers that

- relate to the field of systems engineering;
- represent new, previously unpublished work;
- advance the state of knowledge of the field; and
- conform to a high standard of scholarly presentation.

Editorial selection of works for publication will be made based on content, without regard to the stature of the authors. Selections will include a wide variety of international works, recognizing and supporting the essential breadth and universality of the field. Final selection of papers for publication, and the form of publication, shall rest with the editor.

Submission of quality papers for review is strongly encouraged. The review process is estimated to take three months, occasionally longer for hard-copy manuscript.

Systems Engineering operates an online submission and peer review system that allows authors to submit articles online and track their progress, throughout the peer-review process, via a web interface. All papers submitted to *Systems Engineering*, including revisions or resubmissions of prior manuscripts, must be made through the online system. Contributions sent through regular mail on paper or emails with attachments will not be reviewed or acknowledged.

All manuscripts must be submitted online to *Systems Engineering* at ScholarOne Manuscripts, located at:

https://mc.manuscriptcentral.com/SYS

Full instructions and support are available on the site, and a user ID and password can be obtained on the first visit.

Uncertainty Quantification (UQ) in Complex System of Systems (SoS) Modeling and Simulation (M&S) Environments

Joseph Marvin, Thomas Whalen, Brad Morantz, Ray Deiotte, and **Robert K. Garrett, Jr.**

■ **ABSTRACT**
Prevailing modeling and simulation (M&S) techniques have struggled to provide meaningful quantitative results in M&S of complex system of systems (SoSs) in the face of an environment filled with complex interacting uncertainties. This paper reports on systems thinking applied to "how" M&S techniques should shift to allow a next generation of quantitative tools and techniques. The imperative is to provide quantitative performance results across the constituent interfaces in a modeled architecture. A five-step statistical and parametric algorithm tool that addresses uncertainty quantification (UQ) is presented. [Improving the utility of UQ data evaluation] A quantitative approach to managing complex uncertainties across modeled interfaces using graph theory is proposed. A future vision for SoS engineering (SoSE) that uses graph theory-based modeling is suggested to improve the utility of tools such as UQ is suggested.

INTRODUCTION

As systems have become more complex, the modeling and simulation (M&S) of these systems and SoS becomes equally, if not more, complex. The usual M&S challenges of performance, fidelity and approximations are joined with new challenges such as emergent behavior influenced by aleatoric and epistemic uncertainty. Understanding this uncertainty in models inherently requires an analysis of the data across the entire model or subcomponents of the model. In many cases, complex SoS models are created from individual system component models. In this situation, aggregated SoS models can introduce inconsistencies in data across component system model interfaces. The authors suggest these data inconsistencies in SoS models can be minimized by the use of graph theory. Once these data inconsistencies are reduced or eliminated through graph theory based modeling, the utility and effectiveness of UQ tools is enhanced. We believe UQ in complex SoS M&S environments is best served through graph theory modeling. This enables better treatment of the data across M&S interfaces and improves the utility of UQ tools and algorithms.

The domain of SoS, and systems of systems engineering (SoSE) have received much attention especially from the United States Department of Defense (DoD). Initial attention was placed on defining what is a SoS, that is when is a collection of systems a SoS. Subsequent efforts focused on defining what SoSE is and the establishment of SoSE communities of practice (CoP). Much of this SoS work initially focused on the notion of extending classic systems engineering processes that have been successfully used for engineering complex systems to the SoSE domain. Through review of the literature and our experiences as engineering practitioners, the authors contend that specific and quantitative tools, techniques, and procedures, to define "how to" implement SoSE is missing.

The decision-maker desperately needs quantitative information from the SoSE about SoS mission goals vs. mission capability at the enterprise architecture level throughout the lifecycle. Our research and systems thinking leads us to believe that graph theory-based approaches facilitate answer this need for quantitative SoSE. Garrett and Deiotte (2013) introduced a new methodology for capturing the essence of the physical SoS, the functions executed within the SoS, the interactions between components (at both the functional and behavioral levels) and the causal nature of the SoS based on an employment strategy. Leveraging graph theoretics, a mission level system of systems engineering (MLSoSE) approach provides an abstract, quantifiable method to assess the nature and quality of a SoS model, addressing sensitivity, uncertainty, and quality of the composition.

The foundations of graph theory spans multiple centuries back to Leonhard Euler's paper published in 1736 entitled Seven Bridges of Konigsberg (Euler 1741). This paper related a problem of a traveling salesman who could only cross each bridge in the city of Konigsberg one time while minimizing his total travelled distance. Euler's formula relating properties of the graph (nodes, edges, and faces) is said to be the origin of topology. Graphs are prolific and can be seen in all facets of science, math, nature, and society. Their application to the SoS problem is bourgeoning as can be seen in the multiple graph representations of SysML, DoDAF, and UML, to name a few. Authors like Garret and Anderson, et al. (2011); Marchette (2010); and Luna, Lopes, Tao, Zapata, and Pineda (2013) have all applied the basic concepts of graph theory to multiple aspects of the SoS problem space.

From architecture and design to analysis and verification and validation, graphs can be shown to be at the heart of the SoS paradigm.

The advantages of utilizing graphs for the representation, exploration, and exploitation of SoSs are innumerable. Because of the flexible nature of graphs and their multiple incarnations (undirected, directed, acyclical, cyclical, etc.) graphs can be utilized to represent static and dynamic characteristics of the SoS while utilizing a standard mathematical formulation for quantitative analysis of the structure. By associating metadata to edges and nodes in the graph and transforming from one graph type to another allows engineers, architects, and analysts alike to explore models of SoSs freely extracting pertinent information regarding behavior, performance, uncertainty, and risks.

Additionally, there has been a large move in the past decade to utilize graphs as large, complex data stores to overcome some of the natural issues with traditional relational databases. This includes being able to associate multiple forms of information with multiple types of relationships thereby allowing the extraction of information from data in a manner that is intuitive and frank. Leveraging a graph-based paradigm for SoS modeling and data bases allows interoperability far beyond the current state of the art.

The authors have had a unique opportunity to collaborate on this paradigm shift needed to enable and empower the SoSE to deliver a new level of M&S to address SoSs. We have discussed this issue with system thinkers at the INCOSE 2013 International Symposium in a research session of the Systems of Systems Working group (SoSWG). We have reached out across domains, authors, and expert practitioners. As a result, we are convinced this shift has the potential for a transformation in SoSE by getting to "how." Efforts are underway to complete a graph-theory modeling approach which incorporates graph databases and multiple open-source tools. Uncertainty and uncertainty quantification plays a central role in a new horizon for SoSE. The next section describes work accomplished on an uncertainty characterization and quantification prototype targeted as a key component in near-term SoSE analyses using graph-theory based modeling techniques.

UNCERTAINTY CHARACTERIZATION AND QUANTIFICATION

Uncertainty characterization and quantification (UCQ) is designed to benefit a client with an SoS model whose uncertainty they wish to characterize and quantify to support evaluation, management, and/or improvement of the M&S or perhaps integrating it into a multi-model synthesis. Every SoS model is based on a set of exogenous assumptions about initial conditions, environmental factors, etc. Uncertainty in the output of the SoS model is a consequence of uncertainty in the values of one or more of these exogenous assumptions. Following (*Mathematical Science Foundations of Validation, Verification, and Uncertainty Quantification* 2012) we refer to these uncertain exogenous assumptions as "parameters." The prototype implementation includes a "toy problem" to take the place of a client's M&S. While much simpler than any real M&S system, the "toy problem" includes nonlinearity and discontinuity to ensure that UCQ for characterization and quantification of uncertainty will be effective in an environment where both the real world SOS and the M&S systems modeling it have strong digital switching elements rather than being purely analog systems that can be modeled by differential equations.

Uncertainty is fundamentally the potential to generate surprises, especially unpleasant surprises. At this stage in its development, UCQ focuses on one single number output from the SoS model for each unique set of parameter values the SoS model is given. We refer to this value as the quantity of interest, or QoI (*Mathematical Science Foundations of Validation, Verification, and Uncertainty Quantification* 2012). A commonly used QoI to assess the performance of a simulated missile defense system is margin (*Mathematical Science Foundations of Validation, Verification, and Uncertainty Quantification* 2012). An interceptor missile is assumed to have a kill radius such that, if the interceptor and the enemy attacking missile come within this kill radius, the attacking missile is destroyed with a positive margin equal to the kill radius, minus the distance between the two missiles. If, unfortunately, interception does not occur, the margin is then the smallest distance achieved between the two missiles, minus the kill radius.

For some of the parameters whose value is uncertain, the client may have evidence to support the belief that the real-life value of the parameter will come from a specified probability distribution. In this case we refer to the parameter as aleatoric. Examples include the failure probability for individual components, the probability of a common mode failure, and the prior probabilities and likelihoods that enter into a Bayesian analysis. Some authors treat the latter two categories as epistemic despite the presence of a well-defined probability distribution.

On the other hand, there may be other parameters whose value is uncertain and nonrandom; all the user has available is a range of possibilities with no evidence to support one possibility in favor of another. Parameters subject to this kind of nonrandom uncertainty are referred to as epistemic parameters. One clear example would be the action of an enemy who makes unpredictable choices, not at random, but for reasons of his own which are not fully known to us.

We treat the client's SoS model as a black box; we do not use any classified or proprietary information about the internals of the M&S. In order to use our system, the client permits a link between it and his own SoS model, enabling UCQ to submit a specific value for each parameter, initiate a run of the simulation model, and receive back a value of the QoI. The client must also specify the range of possible values for each epistemic parameter and the probability distribution function for each aleatoric parameter. Figure 1 provides the context of UCQ acting on the SoS model (M&S system) toy problem.

Using this information, UCQ first creates a representative random sample using the Latin Hypercube sampling method (McKay and Conwer 1979). Using a geometric analogy, we refer to the vector specifying a unique value for each uncertain parameter as a point in a multidimensional parameter space.

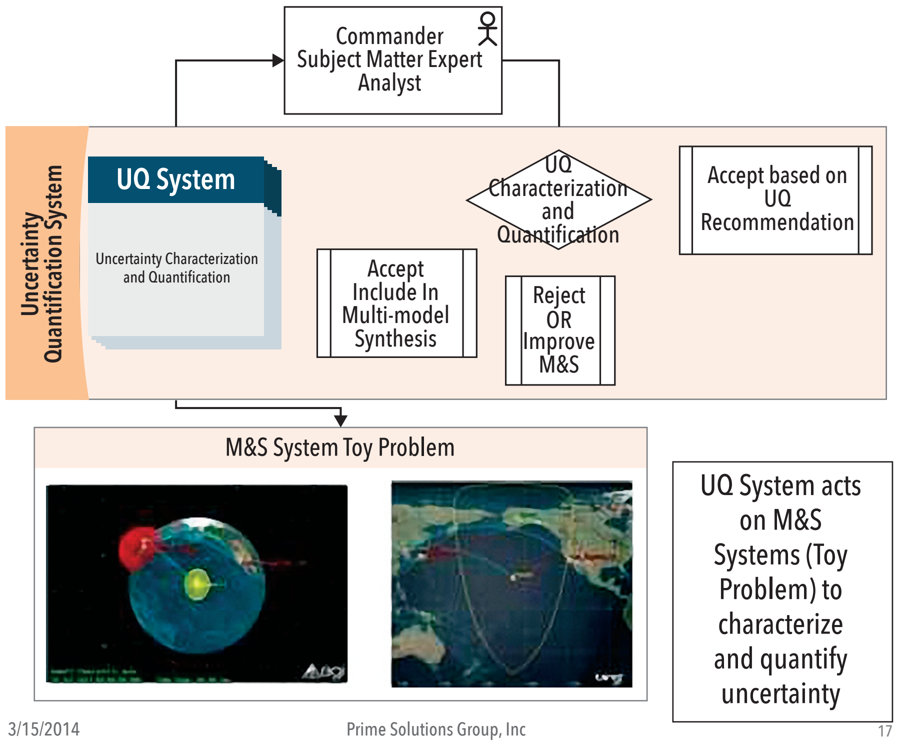


Figure 1. UCQ context in SoS model evaluation

In order to treat aleatoric and epistemic parameters consistently without discarding legitimate probability information about aleatoric parameters, or introducing spurious probability assumptions about epistemic parameters, we apply a generalized probit transformation (Bliss 1934; Cleophas and Zwinderman 2012) to replace each possible value that a given aleatoric parameter might take on with its cumulative probability in the unit interval. We also use a simple linear transformation to map the range of possible values for each epistemic parameter into the unit interval. This maps the higher dimensional space, defined by parameters expressed in natural units, to a unit hypercube. For reasons that will become apparent shortly, we refer to this unit hypercube as "batspace."

For each point in the parameter space selected by the random sampling process, UCQ queries the client's SoS model, which takes those specific parameter values, runs the simulation from start to finish, and returns the value of QoI. UCQ uses this representative random sample to output a standard statistical report of mean, standard deviation, skewness, kurtosis, and other statistics giving information about the global behavior of the QoI, as parameter values vary according to the aleatoric or epistemic information available.

Following the conventional statistical analysis, we institute an exploited search procedure in which biologically inspired agents (Brownlee 2012) ("bats") seek out areas of noteworthy performance (ANPs) (Schultz et al. 1992. ANPs come in three varieties. An upside ANP is an area in parameter space characterized by exceptionally high values of QoI; a downside ANP is characterized by exceptionally low values; and an unstable ANP is characterized by rapidly changing QoI values associated with small changes in the values of the parameters.

An ANP is a hazard if it indicates the mere possibility of failure; it is a risk if failure has a significant degree of plausibility or probability (NRC 2009).

The database formed by aggregating all of the upside, downside, and unstable ANPs forms the basis for a global characterization and quantification of uncertainty. Dempster's (1967) upper probability is a mathematical tool well suited for combining aleatoric and epistemic uncertainty without loss of information or introduction of spurious assumptions. Our approach normalizes the result of this calculation to create the membership function of an evidence-based fuzzy set of plausible values of QoI. The plausibility of failure is the highest membership of any negative value of QoI in the fuzzy set of plausible values. The probability of failure is always less than or equal to the plausibility of failure; probability cannot be precisely quantified in the presence of epistemic uncertainty. The total uncertainty of the SoS model is quantified by the area under the membership function, the "sigma count." The global characterization and quantification of uncertainty, summarized by the two key scalar metrics quantity of uncertainty and plausibility of failure, is useful for comparative assessment of one SoS model against competing SoS models of the same real world scenario.

We created a toy problem representation of an SoS model. The toy problem is a simple missile defense scenario that mimics the effects of aleatoric and epistemic parameters. These effects include a "kill radius" and a hypothetical "Hawaii" effect. These terms appear throughout the following explanation of the five major steps in our technical feasibility demonstration.

Step One: Consistent Treatment of Aleatoric and Epistemic Parameters. In order to treat aleatoric and epistemic parameters defined on a variety of scales of measurement in a consistent manner without losing valuable probabilistic information about the aleatoric parameters, and also without introducing spurious probabilistic assumptions about the epistemic parameters (Aven 2010), we transform the parameter vectors defined in natural units in parameter space into a multidimensional unit hypercube, referred to as batspace.

Biologically inspired exploited search agents, "bats," operate in batspace to find ANPs which, when translated back to natural units in parameter space and submitted to the client's SoS model, generate areas of performance noteworthy for exceptionally high, exceptionally low, or exceptionally unstable performance.

The mapping process for an epistemic parameter is very simple. A given specific value of such a parameter in natural units is transformed into a value in the unit interval by subtracting the lower bound of the parameter's range and dividing by the width of that range. The reverse transformation is simply the mathematical inverse of that process.

The transformation of an aleatoric variable from parameter space to batspace is more complex. It uses a probit transformation (Bliss 1934) to hold in abeyance the probabilistic information about the aleatoric variable in such a way that it can be ignored by the bats operating in batspace but restored when batspace values are transformed back into natural units in parameter space. In UCQ we transform a value in parameter space by using the probit transformation giving its cumulative probability, and we transform a corresponding point in batspace using the inverse probit transformation to find the value in natural units whose cumulative probability is that given value in the unit interval. Since the literature on probit transformation is generally focused on

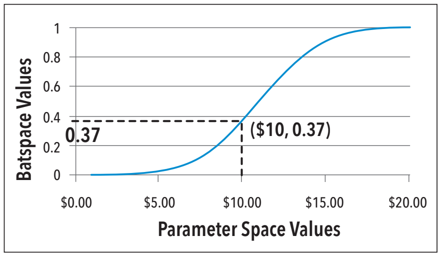


Figure 2. Cumulative probability graph

Gaussian distributions, we use the phrase "generalized probit transformation."

These transformations can be visualized using an ordinary cumulative probability graph. Suppose that one of the parameters of an M&S is an amount of money that is normally distributed with a mean of $11 and a standard deviation of $3. If we needed to find the location in the unit interval in batspace corresponding to a dollar amount of $10, we can see it in Figure 2 by looking vertically upward from $10 on the parameter space axis and reading the value of the corresponding element of batspace as 0.37. Contrariwise, if we had a value of 0.37 in batspace and needed to find the corresponding value in natural units in parameter space, we would look horizontally from 0.37 on the vertical batspace axis and read the corresponding value of $10 in parameter space.

Since the toy problem that we used to demonstrate the performance of the prototype system for characterization and quantification of M&S uncertainty has six uncertain parameters, the parameter space and corresponding batspace in the present prototype each have six dimensions, with a seventh dimension representing the QoI. Increase in the number of dimensions is straightforward and adds only linearly to the amount of computation required.

Step Two: Standard Statistical Characterization and Quantification of Uncertainty. In this step we treat each transformed parameter in batspace as a uniformly distributed random variable. This is standard practice in Monte Carlo simulation for aleatoric uncertainties. If the transformed value in the unit interval is given a uniform distribution, the inverse transformation will give the correct probability distribution in natural units in parameter space. For epistemic uncertainties, the principle of insufficient information (Aven 2010) is frequently used to justify the imposition of a uniform distribution. These distributional assumptions in batspace are only temporary for this stage of the analysis and will be discarded in the following stage; however, they are standard practice and the statistical results of this assumption are provided to the client as a familiar set of tools for assessing SoS models.

The Latin Hypercube sampling approach (McKay et al. 1979) is a widely used method to generate a representative random sample of inputs to be used in assessing a computer simulation system. We use this method to perform a stratified random sample on the assumption of a multivariate uniform distribution discussed above.

Step Three: Search for Areas of Noteworthy Performance. The next step is to assign each sample point in the representative sample described above to a biologically inspired parallel exploited search agent. These search agents are loosely modeled as bats that use active sonar at short range to detect tiny mosquitoes by emitting chirps at varying frequencies depending on the local concentration of prey. This variation allows the bats to also use passive sonar (Bohn 1838) at longer ranges to see which other bats in their family are achieving better hunting success and fly towards those more fortunate family members on the basis of two instincts. One instinct is to approach more successful family members, and the other is to fly toward family members who are nearby. Each bat's decision-making is done in parallel, independently of the decision-making process of any other bat relying only on fellow family members' location and success.

Since there are three classes of ANPs: upside, downside, and sensitivity, there are effectively three different species of bat. The concentration of the prey sought by upside bats at any point in batspace is directly proportional to the value of QoI in the corresponding point in parameter space. Downside bats hunt for a prey whose concentration in batspace is negatively proportional to the value of QoI in parameter space. The behavior of our sensitivity bats is least comparable to that of any real bat. These bats are motivated to fly towards nearby family members whose success rate is maximally different from theirs, whether greater or lesser.

When the bats have arrived at their next location, their position in batspace is transformed back to natural units as discussed above. This vector of parameter values is transmitted to the client's SoS model which performs a simulation run and returns the QoI. Once every bat has received the QoI value corresponding to its new location, they determine which family member to fly towards and move to a third location and the cycle continues. The current implementation of the working prototype converges within 10 or fewer iterations, except in the more complicated search for sensitivity ANPs. The prototype has 30 bats divided into five families of six bats each; since all three species use this structure there are effectively 15 families, 90 bats in all. It is a straightforward matter to increase the number of bats, with only a linear increase in the total amount of complexity required. The result of the collection of points visited by the bats during the search is a sample which is divided into separate ANPs, one for each family for each of the three types of ANPs sought. The entire sample is used in step four, while the data from each of the 15 families is used separately in Step Five.

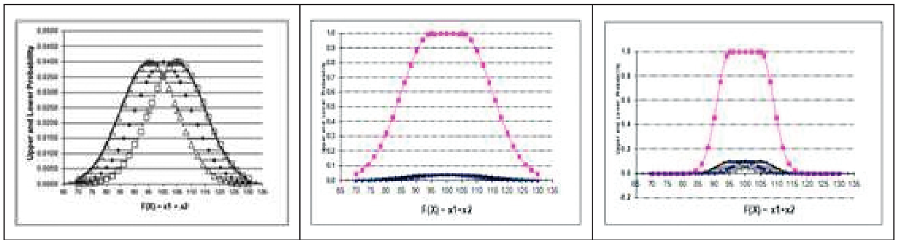
Step Four: Global Characterization and Quantification of Uncertainty. The ideal real-world system would be one that delivered excellent performance regardless of what the values of the possibilistic and probabilistic parameters happened to be. An uncertain system is one that works well in some circumstances and poorly in others. Uncertainty in a purely probabilistic or aleatory system is defined by the variance, probability of failure, and other statistical dispersion measures of the QoI across the probabilistic variations of the probabilistic inputs. Contrariwise, when all the dimensions are possibilistic, simple interval analysis (Moore et al. 1996) suffices. For the general situation of UQ for simulation models of highly complex real-world systems, neither of these standard approaches is entirely satisfactory, as discussed above. Consider an extremely simple system with just one probabilistic variable x1 which is normally distributed with mean 100 and standard deviation 10, and one possibilistic variable x2 ranging from -5 to +5 with no evidence to support any particular probability distribution. The response surface for this extremely simplified example is F(X) = x1 + x2.

In Figure 3, Graph A, the solid diamonds represent the probability density function of F(X) when x2 is equal to 0. The triangles represent the probability density function of F(X) when x2 is equal to -5. Note that pessimistic values of F(X) are more probable when x2 is negative and optimistic values of F(X) are more probable when x2 is positive.

The solid black curve is the maximum probability density of F(X) over all values of x2 for a given value of x1. It was called the upper probability by Dempster.

In 1967 Dempster proved that the probability of a partially specified event must be less than or equal to the upper probability in a mathematically defined class of situations that includes mixed models with possibilistic and probabilistic variables.

The upper probability curve can be approximated from the collection of values generated by the bats. Conceptually, we can graph each observed value of F(X) against the upper probability of the most plausible point in input space observed to generate that value of F(X), then connect the dots



Graph A: Derivation of Plausibility | Graph B: Set of Plausible Values, s.d. = 10 | Graph C: Set of Plausible Values, s.d. = 4

Figure 3. Plausibility

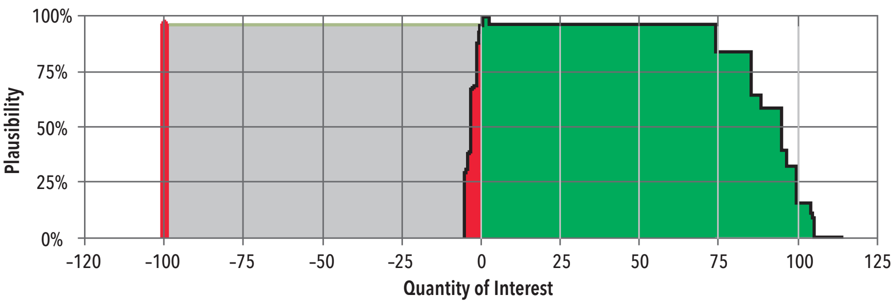


Figure 4. Evidence-based fuzzy set of plausible values (with Hawaii effect)

Table 1. Bat Family 1 sensitivity search

	x1	x2	x3	x4	x5	x6	QoI
mean	80	23	103	56	46	103	-22
median	81	16	100	59	52	100	14
maximum	100	100	141	100	88	141	125
minimum	14	0	94	0	0	98	-100

to draw the convex hull and a smoothed version of it.

The area under an upper probability curve of the extremely simple function F(X) = x1+x2 is proportional to the range of the possibilistic variable but is insensitive to the variance of the probabilistic variable; similar difficulties apply to more realistic cases. Calculating a normalized plausibility by dividing each upper probability by the maximum plausibility can be used to define a fuzzy set of plausible values. The area under this curve, known as the sigma count, is positively related both to the variability of probabilistic variables and the range of possibilistic variables. This area constitutes our measure to quantify the global uncertainty of a simulation model for assessing highly complex real-world systems. The global uncertainty of the UQ system's analysis of a given simulation model is represented by the sigma count of the plausibility membership function, which is just the area under the curve. Since it is the integral of a curve found by dividing one probability density by another, it is dimensionless but serves to compare the total uncertainty of one model with another and, by extension, the total uncertainty of the two real-world systems being so modeled.

In a totally certain system, the graph would be a vertical line at the certain QoI value. In a totally uncertain system, the graph would be a horizontal line from −infinity to +infinity and uncertainty would be infinite.

The plausibility of failure is the plausibility of the most plausible negative QoI; graphically it is the maximum height of the red area.

Figure 3, Graph B, shows the fuzzy set corresponding to the example discussed above, while Figure 2, Graph C, shows the fuzzy set arising from the same example, except with a standard deviation of only 4 for x1.

Figure 4 shows the global quantification of uncertainty for the M&S with the Hawaii effect included. The gray area represents values of QoI which are not observed, but which contribute to the total uncertainty because of the penalty value of -100. The total quantity of uncertainty for this M&S is 184.94 and the plausibility of failure is 96.14%.

Step Five: Local Characterization and Quantification of Uncertainty. Global characterization and quantification of uncertainty is useful for comparing the overall effectiveness of competing variations of the real world SoS (Volkert et al. 2013). However, more is needed in order to achieve the full potential of characterizing and quantifying the uncertainty in an SoS model. Step 5 introduces additional statistical analysis steps that help to graphically characterize and quantify uncertainty. Understanding what goes on at each individual ANP is necessary to inform the process of identifying errors and inadequacies in the simulation system in order to improve the simulation and/or to recognize opportunities to design a synergistic multi-model M&S that exploits the specific strengths and overcomes the specific weaknesses of complementary individual SoS models. Moreover, it is through information about the various ANPs of an SoS model that the client is equipped to discover the strengths and weaknesses of the real world SoS. The following sections describe mathematical, statistical, and graphical characterization and quantification of local uncertainty in the particular area whose performance is noteworthy for high, low, or rapidly changing values of QoI.

Mathematical Characterization and Quantification of Local Uncertainty. The local mathematical information begins with a table showing the location in parameter space of the mean values of all the parameters and of QoI for all bat families of a particular type: upside, downside, or sensitivity. Following this there is a series of reports for each individual ANPs.

The first element of this series of reports is a table giving the mean, median, maximum, and minimum values of each parameter in the ANP in question, together with the mean, median, maximum, and minimum values of the QoI in the ANP. We will use the results from Bat Family 1 in the above sensitivity search as our example shown in Table 1.

Statistical Characterization and Quantification of Local Uncertainty. The statistical portion of the information provided for each individual ANP begins with a stepwise quadratic regression equation fitted to the points visited by one particular bat family, with the exclusion of the first two sets of points. The resulting quadratic form potentially including all of the parameters is a meta-model of the behavior of the SoS model focused on just the neighborhood of the ANPs inhabited by that particular bat family.

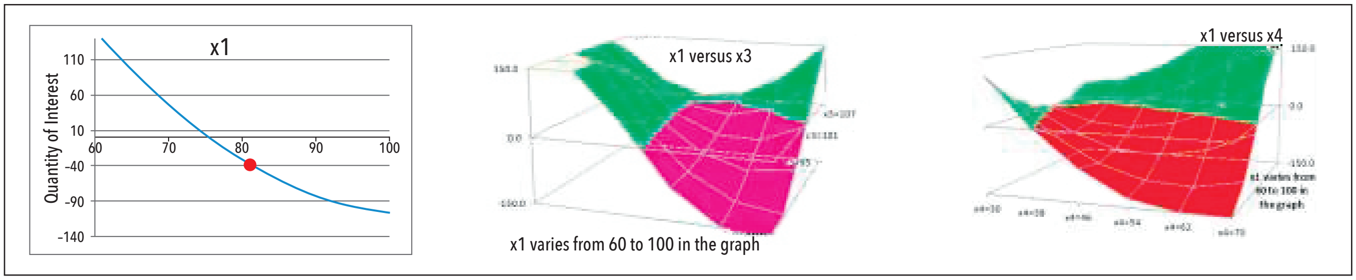


Figure 5. Graphical characterization and quantification of example ANP

A variety of statistical analysis tools can be applied to the data collected for each ANP. In our current approach a local quadratic regression for each combination of three parameters vs. QoI and reports which combinations give the best statistical predictive power.

Graphical Uncertainty Characterization and Quantification of Local Uncertainty. The three parameters with the highest collective predictive power in a local quadratic approximation for this ANP are x1, enemy aim point 1; x3, kill radius 1; and x4, attacker aim point 2. The graphical characterization and quantification of an ANP begins with an array of nine graphs. Each of the three parameters in the most predictive set of three has one row and one column in the table. The median value of the predictor in question within the given ANP is indicated by a red dot on the quadratic regression line.

Figure 5 gives the graphical characterization and quantification of uncertainty for the example ANP discussed above. Red regions in the three-dimensional space portrayed in these graphs indicate that the QoI is negative, meaning that the M&S has predicted failure of the real world SoS under simulation. Green regions indicate positive values of QoI.

The first graph approximates the QoI as a local univariate quadratic function of x1, enemy aim point 1. Since the median value x1 is already above the corresponding defending aim point x2 in the cluster found by Bat Family 1, increases within the boundaries of the neighborhood of this ANP consistently lead to more favorable values of QoI and vice versa.

The three-dimensional graph in the center, showing the local quadratic approximation of QoI as a function of x1 and x3 contains an important interaction effect. Within the local area of this ANP, performance is negative only when X1 and X3 are both low. This can be seen in the lower front right corner of the 3D graph. If X1 is high, that means in the context of this local ANP that it is close to X2; the two missiles are close together, so that a value of X3, the kill radius, within the local area of this ANP can still lead to positive QoI. Similarly, if X3 is high, the kill radius is wide enough that x1 can be in the low end of its range within this local area, relatively far away from the intercepting missile assigned to it, and still be destroyed, resulting in a positive value of QoI.

The third three-dimensional graph shows the interaction between the first enemy aim point x1 and the second enemy aim point x4 in the area inhabited by Bat Family 1. The red color is seen for high values of x1; for x4, the graph is green for values close to 50 and red for values either much higher or much lower than 50.

The second component of local graphical characterization and quantification of uncertainty consists of a graph comparable to the one for global characterization and quantification of uncertainty, with QoI on the horizontal axis and plausibility on the vertical axis. The amount of uncertainty indicated by the gray area is much larger than the area of uncertainty in the global characterization and quantification of uncertainty for the toy problem that includes the Hawaii effect. The same possibility of a penalty of -100 with one hundred percent plausibility exists in both cases, but in this local ANP, which was found by bats specifically seeking maximum sensitivity, the green area is smaller and concentrated on the most favorable end. The above example is shown in Figure 6.

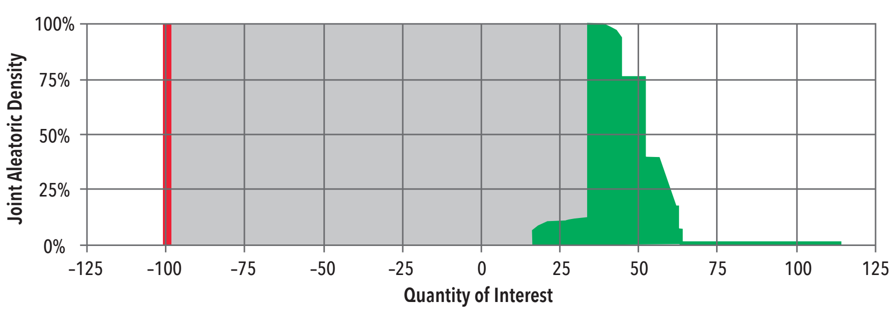


Figure 6. Plausible instability from sensitivity ANP 1

SUMMARY

Uncertainty Characterization and Quantification. The goals of this project are to demonstrate feasibility of a system that will characterize and quantify the output uncertainty produced by an M&S arising from a combination of aleatoric and epistemic input uncertainties. UCQ system performs this characterization and quantification both globally and locally in terms of specific ANPs. Additionally, UCQ is designed to be able to be scaled up to high dimensionality in order to accommodate large models in a timely basis without requiring a supercomputer. Having such a tool will make it possible to evaluate and compare M&S SoS, for both evaluation and improvement as well as possible synthesis into a multi-model approach. Graph data bases as implemented by graph theory facilitate UCQ analysis of the data. Our toy problem developed in Microsoft Excel follows this concept by passing the actual data from one component to the next.

Our vision for UCQ is to build a web service open architecture prototype that facilitates coupling and interfacing with any number of modeling and simulation tools. A first integration in the next 24 months will be with a graph theory-based architecture modeling tool. Our open architecture approach allows a service and application layer UCQ capability current complex M&S environment, even those not built on service oriented constructs. The substantial shift to graph theory-based M&S in the COP will require an ability to demonstrate the value of graph theory results compared to legacy SoS models. UCQ will serve as an initial independent variable.

Graph Theory SoS M&S. Complex SoS M&S is a fertile, undefined playing field. To the authors there is a critical need for multiple layers of both modeling and simulation that span the conceptual to the continuous. Due to the nature of the SoS

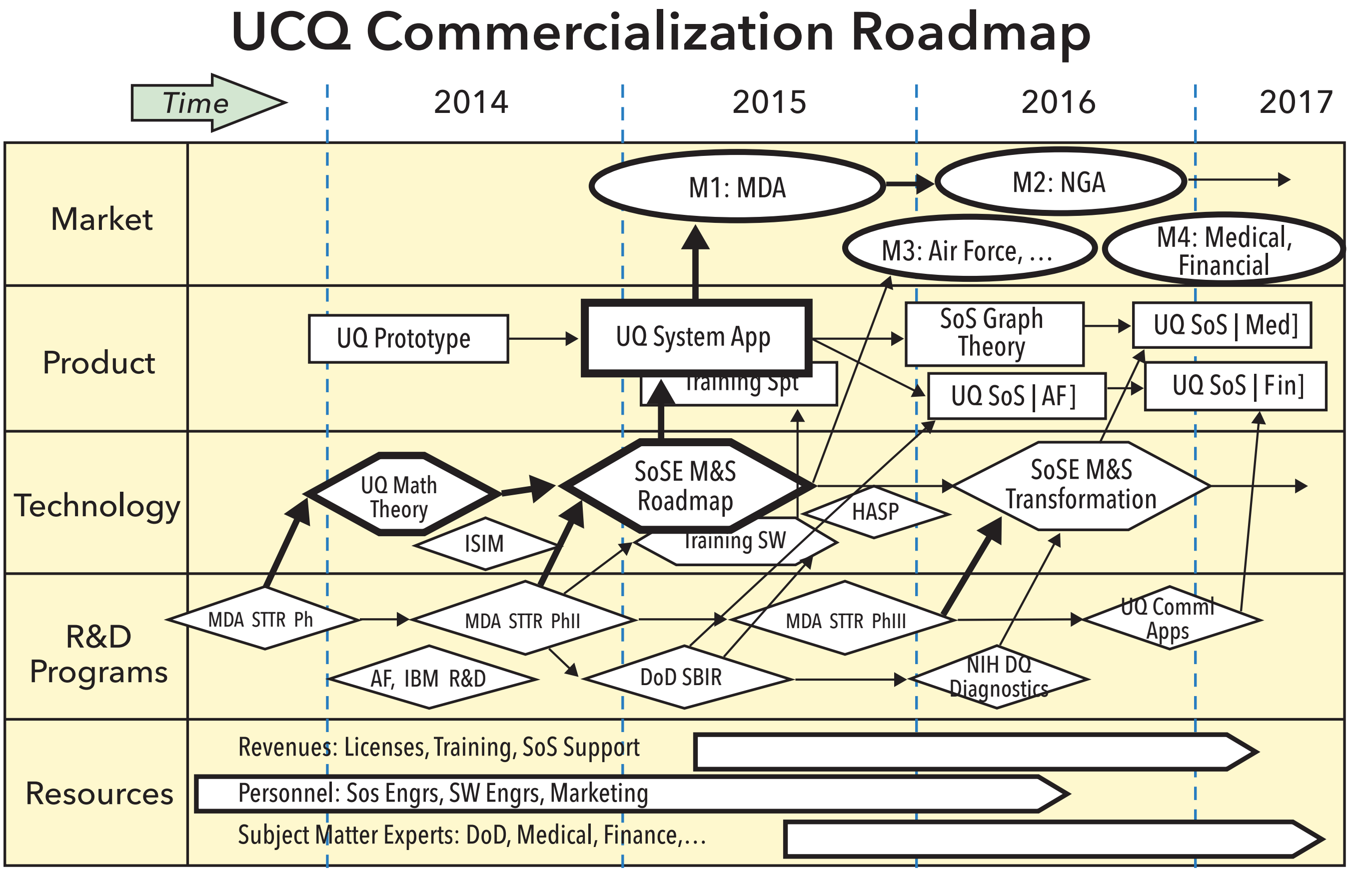


Figure 7. UCQ commercialization roadmap

(their sizes and complexities) a layered approach is necessary to be able to capture both the global and local characteristics of the SoS. System graphs and subgraphs using graph theory provide this approach. Our vision is one of tiered, integrated M&S tools that have the ability to perform UCQ analysis at global and local levels which can feed into design-level M&S and testing and ultimately to verification, validation and assessment of the SoI within one seamless process. The ability to leverage common data structures and mathematics will determine the community's ability to achieve this goal. We propose the migration of all tools, techniques and procedures to a graph basis to allow this transformation and to fully reap the benefits of SoS models.

Future of SoSE and Roadmap. A confederation of small business innovative research contractors led by government change agents has taken on the challenge to demonstrate the value of the approach described in this paper. We view INCOSE as a key supporter of the systems thinking necessary to shift M&S practices that achieve benefits described in this paper. Our reach is expanding rapidly, and we find interest in many adjacent domains and enterprises. We are on a roadmap that builds a COP and conducts a series of technical interchange meetings (TIMs) with invited stakeholders. The goal of these TIMs will be to advance ideas, applications, and demonstrate results of "how" SoSE will conquer complex SoS M&S through UCQ. The commercialization roadmap we are currently following is provided as Figure 7.

The authors expect to continue working this roadmap in a variety of forums including scientific technology transfer (STTR) projects and the INCOSE SoS WG. ■

REFERENCES

- Aven, T. 2010. "On the need for restricting the probabilistic analysis in risk assessments to variability." *Risk Analysis: An International Journal* 30, no. 3: 354-360
- Bliss, C. I. 1934. "The method of probits." *Science* 79, no. 2037: 38-39.
- Bordawekar R., B. Blainey, C. Apte, and M. McRoberts. 2011. "Analyzing analytics, part 1: A survey of business analytics models and algorithms."
- Brownlee, J. 2011. *Clever algorithms: nature-inspired programming recipes*. Jason Brownlee, 2011.
- Cleophas, T. J., and A. H. Zwinderman. 2012. "More on nonlinear relationships, splines." Statistics Applied to Clinical Studies, 277-288.
- Deiotte, R., and R. K. Garrett Jr. 2013. "A Novel Approach to Mission-Level Engineering of Complex Systems of Systems; Addressing Integration and Interoperability Shortfalls by Interrogating the Interstitials, Missile Defense Agency."
- Dempster, Arthur P. 2008. "Upper and lower probabilities induced by a multivalued mapping." In *Classic works of the Dempster-Shafer theory of belief functions*, pp. 57-72. Berlin, Heidelberg, DE: Springer Berlin Heidelberg.
- Garrett Jr, R. K., S. Anderson, N. T. Baron, and J. D. Moreland Jr. 2011. "Managing the interstitials, a system of systems framework suited for the ballistic missile defense system." *Systems Engineering* 14, no. 1: 87-109.

- Helton, J. C. 2009. *Conceptual and computational basis for the quantification of margins and uncertainty*. No. SAND—2009-3055. Sandia National Laboratories (United States). Funding organization: US Department of Energy (United States).
- Luna, S., A. Lopes, H. Yan See Tao, F. Zapata, and R. Pineda. 2013. "Integration, verification, validation, test, and evaluation (IVVT&E) framework for system of systems (SoS)." *Procedia Computer Science* 20: 298-305.
- Marchette, D. 2010. "Class cover catch digraphs." Wiley Interdisciplinary Reviews: Computational Statistics 2, no. 2: 171-177.
- Martinez, J., R. Carroll, J. Morris, and K. Bohn. 2008. "A Study of Mexican Free-Tailed Bat Chirp Syllables: Bayesian Functional Mixed Models for Nonstationary Acoustic Time Series." *The Journal of the Acoustical Society of America*, Vol 124: 1838.
- Mathematical Science Foundations of Validation, Verification, and Uncertainty Quantification. March 28, 2012. http://sites.nationalacademies.org/DEPS/BMSA/index.htm (accessed 2013).
- McKay, M. D., R. J. Beckman, and W. J. Conover. 2000. "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code." *Technometrics* 42, no. 1: 55-61.
- Moore, R. E., R. B. Kearfott, and M. J. Cloud. *Introduction to interval analysis*. Society for Industrial and Applied Mathematics.
- "National Research Council.2009. "Science and Decisions: Advancing Risk Assessment. Washington, US-DC: The National Academies Press.
- "National Research Council. Making Sense of Ballistic Missile Defense: An Assessment of Concepts and Systems for U.S. Boost-Phase Missile Defense in Comparison to Other Alternatives." 2012. Washington, US-DC: The National Academies Press.
- Raiffa, H. 1968. *Decision Analysis: Introductory Lectures on Choices und Uncertainty*. Addison-Wesley.
- Schultz, A. C., J. J. Grefenstette, and K. A. De Jong. 1995. "Applying genetic algorithms to the testing of intelligent controllers." *Applying Machine Learning in Practice, IMLC-95*.
- Volkert, Richard, Jerrell Stracener, and Junfang Yu. "Incorporating a measure of uncertainty into systems of systems development performance measures." *Systems Engineering* 17, no. 3 (2014): 297-312..
- Whalen, T. , B. Morantz, and M. Cohen. 2013. "Uncertainty Quantification for Possibilistic/Probabilistic Simulation." IFSA World Congress and NAFIPS Annual Meeting (IFSA/NAFIPS).

ABOUT THE AUTHORS

[Editor: Author biographies were current when the paper was initially published in 2014.]

Joseph Marvin is an INCOSE Expert Systems Engineering Professional (ESEP) and chair of the INCOSE VSME WG. Mr. Marvin is president of Prime Solutions Group, Incorporated (PSG) a systems engineering, software development, and innovative research VSE. His background is in defense systems engineering with over 30 years of systems engineering in major DoD space programs in government and industry. He has led PSG to certification as an IBM business partner in systems and software acceleration using IBM Rational Collaborative Lifecycle Management (CLM) tools and techniques. He leads the PSG small business innovative research team in system of systems (SoS) uncertainty quantification.

Brad Morantz is a senior member and vice chair IEEE Phoenix Computer Society. He is also a member of Computational Intelligence Society and Association for the Advancement of Artificial Intelligence (AAAI). His development and practice background in defense programs include M&S, fault analysis, and decision elements of the Kinetic Energy Interceptor program. He has also applied fundamental research in data mining, statistical analysis, and automated image pattern recognition into applications for the Joint Improvised Explosive Device Organization (JIEDO). He holds a PhD in decisions sciences, Georgia State University; MS, decision sciences, Georgia State University; and BS, CIS/EE, Cleveland State University.

Thomas Whalen, PhD, is director of research, Frontline Healthcare Workers Safety Foundation. He served as professor of decision sciences, Georgia State University, where he taught modeling, statistics, probability, and advanced subjects. He has significant experience as a senior member, Institute of Electrical and Electronic Engineers (IEEE); associate editor, IEEE Transactions on Fuzzy Systems; and board of directors, North American Fuzzy Information Processing Society. Dr. Whalen is a nationally known expert in decision theory and systems with over 100 papers, handbooks and lectures. He conducts innovative research for the Department of Defense and National Institutes of Health on advanced decision theory and uncertainty.

Ray Deiotte is the senior scientist and director of innovation for the ISSAC Corp. Mr. Deiotte holds a BS in physics from the University of Colorado. He has post graduate studies in physics, software engineering, systems engineering, data mining and management, and cognitive reasoning. Mr. Deiotte worked many years in defense system of systems modeling and simulation, integration, and system analytics and assessment. His current research focuses on integration of graph-based knowledge systems with evolutionary learning algorithms and advanced modeling and simulation for highly interactive predictive systems.

Robert K. Garrett, Jr. earned a BS in materials engineering from Purdue University in 1981 and an MS in materials engineering from Purdue University in 1983. He worked in research and development with the Naval Surface Warfare Center for 27 years. His expertise is in systems of systems engineering, integration of diverse technologies into weapon systems, and the application of materials science and continuum mechanics to weapon systems development. In 2010 Mr. Garrett joined the Missile Defense Agency where he is currently a technical advisor supporting the director for modeling and simulation. He has focused his attention on adapting, developing, and exploiting modeling, simulation, and analysis tools for system of systems engineering. Mr. Garrett was recognized by the International Council on Systems Engineering (INCOSE) for his published work, "Managing the Interstitials: A System of Systems Framework Suited for the Ballistic Missile Defense System". The work was selected by a committee of Associate Editors as the recipient of the 2011 Best Paper Award.

Measuring the Uncertainty Impacts During the Systems Engineering Lifecycle

David Flanigan and **Jeffery Dixon**

ABSTRACT
Uncertainty is a large part of the systems engineering development process. Particularly absent is the quantification of uncertainty of the threat, operating environment, and friendly force factors at each step of this lifecycle. This paper will explore a methodology to quantify the amount of uncertainty and the interdependencies of the uncertainty factors during the development. Included for consideration are internal and external factors and their contribution to the overall system uncertainty. An illustrative example is provided to exercise this methodology.

INTRODUCTION

We are motivated to quantify the uncertainty inherent with the numerous inputs that affect a system development cycle. These inputs may be in the form of uncertainty in the threat capability, the operating environment that the system was designed for, or the system itself through technical performance, tactical implementation, and program acquisition. If uncertainty was not considered, professions such as requirements analysts, concept developers, and testers are in danger of starting development of a system that is not prepared to handle the representative threats or operate in a representative environment. Review of the current literature indicates a general lack of quantification of the total uncertainty and how component uncertainty factors are related to each other. Because of this lack of quantification, we propose a methodology in which to identify and quantify the uncertainty factors, using an illustrative example.

LITERATURE REVIEW

The term uncertainty is reviewed in the literature to understand how it is used in systems development as well as understanding how uncertainty may be quantified. McManus and Hastings (2004) develop a framework to understand how uncertainty may be mitigated during project development, with two significant terms emerging from their research. Their first term identifies a lack of knowledge about the system, where facts are not known, but are required to be collected. Their second term is the lack of system definition, where knowledge about the system is incomplete and not specified. These terms start to form a basis of how one may categorize the types of uncertainty encountered during a systems development effort. They also define several entry points of uncertainty during the program: the system may not work as designed; the system may work but not up to expectations; the market may shift; the need for the system may shift. Their identification of mitigation efforts includes use of performance margins, use of subsystem redundancy, use of program testing, use of general components, use of upgrades, and use of standard interfaces. We may utilize their research as motivation to identifying a particular strategy towards quantifying uncertainty and how it may be managed.

Flage and Aven (2009) perform research on the level of uncertainty intervals as being dependent on where one is in the systems development lifecycle, based on industry practice. Their research identifies +/- 40% uncertainty interval in the feasibility phase, +/- 30% interval in the concept development phase, and +/- 20% interval in the engineering phase. Their findings indicate that the uncertainty will gradually be reduced as the development progresses. During this progression, the level of system detail in the early phases is low, with an increased level of detail as the system reaches the engineering phase. We may use this to identify a means of an expected level of uncertainty as a function of the lifecycle, in which to measure our uncertainty methodology.

Yang et al. (2007) identify additional contributions of uncertainty to a program, specifically for software development: changes in requirements, a general lack of requirements understanding, lack of historical data, and a lack of estimation methodology. In response to the uncertainty levels, they offer a means of agile methods in order to embrace change, promote frequent delivery, and implement a simple design to reduce uncertainty. Averyt et al. (2013) explore the concept of multiple scenarios set in the future, which can identify the available tradespace that may emerge from the decisions made in earlier stages. We can adapt this concept to explore how well the uncertainty can be measured and managed based on the early-stage systems engineering decision.

Boehm (2009) identifies several challenges to Department of Defense (DoD) software programs that will include handling emergent requirements, rapid changes to programs, and incomplete visibility of programs within a system of systems construct. He introduces a "cone of uncertainty" concept that reflects a gradually decreasing level of uncertainty as the system concept matures, that is measured at the initial development phase (CONOPS phase), at the specifications phase, and at the final delivery (Initial Operational Capability). In order for the program to remain within this cone of uncertainty, he requires programs to increase investments, implement competitive prototyping, and conduct concurrent engineering. He claims that failure to evolve the program may introduce a second cone of uncertainty based on uncertainties in competition, technology, and organization. Boehm's cone provides a visual reference in which to reduce the uncertainty levels based on a particular time frame to meet a successful program development.

Utility functions are developed in order to compare dissimilar elements using a common or normalized scale. It is most often used for comparison of alternatives, such as a trade study. Buede (2009) refers to these as value curves, but they provide the same purpose of utility functions. Kossiakoff and Sweet (2011) describe methods for using utility functions to perform trade off analysis. With regard to this research, we develop a series of utility functions to represent general relationships between different subsystems and to quantify the resultant output when provided an input from a transmitting subsystem.

From our literature search, we can conclude several observations. One is that the applications of uncertainty are varied within the hardware and software domains. There are many ways in which to measure the uncertainty of system progress, requiring a formalized means to identify and evaluate the causality between factors that cause an increase in uncertainty. So far the literature does not indicate a means to identify and measure the causality between different uncertainty factors, rather focusing only on a single primary source of uncertainty. Based on the literature review, we are motivated to introduce a methodology in order to measure this causality and the effects of differing uncertainty impacts on the system development. This measurement may provide decision makers with an understanding of what amount of uncertainty exists in a particular development stage and may assist decisions to continue or discontinue the program based on the expected vs. actual amount of uncertainty. By utilizing such a quantitative process, decision makers may also be informed of the level of risk they are undertaking while considering different alternatives or concepts while addressing the problem.

UNCERTAINTY CALCULATION METHODOLOGY

A five-step methodology is developed to quantify how various uncertainty elements can be measured to describe the overall uncertainty to the functions that are needed to successfully execute the system's mission. The methodology is described as: identify the uncertainty areas, develop the uncertainty utility function, describe the uncertainty interdependency, collect the uncertainty inputs, and conduct the overall mission uncertainty analysis.

Step 1: Identify the Uncertainty Areas

This step will start to identify the different uncertainty areas that will influence the mission execution of the system under consideration. These uncertainty areas may be internal, whether it is the subsystem / component development, or external system use by the operators of the system. The subsystem / component development may have their requirements changed in response to another threat or environment that may impact the system's development timeline. The subsystem may be developed and optimized for one particular environment, and therefore be marginalized when expected to perform in another environment. Funding changes may influence the subsystem development that was not foreseen, therefore delaying the capability further in the future than anticipated. Additional tactics, techniques, and procedures (TTP) that were not envisioned in the original design may also influence the system performance. Adding additional systems or constraints onto the original design may also reduce the system efficiency or may induce additional interface and reporting requirements that may reduce performance in the system's operational environment.

These uncertainty areas may also be externally driven with unknowns about how the threat / target (also referred to as "red") will operate, or a change in their technical capabilities that the system was originally designed to address. Other external uncertainties may include changes in the operating environment that the system was designed to perform within. Systems may be designed with an initial expectation of how the threat will progress in time to address the expected level of future threat performance when the system is to become operational. If the threat increases their developmental cycle greater than the system's projections, then an inferior friendly (also referred to as "blue") system will be developed. If the threat changes their tactical employment in an unanticipated direction, then the developed system's performance may not be sufficient to address the threat. If the environment suddenly changes to require the system to operate in a different environment in which it was not designed for, then the system performance may be inferior to its original design. The output of this step is an identification of all internal and external factors that may contribute to the overall system uncertainty.

Step 2: Develop the Uncertainty Utility Function

This step identifies the relationship between the input (uncertainty, whether that comes from an internal or external source), and the resultant output on system / subsystem performance. This relationship draws from utility theory, in which the input and output scale are normalized between multiple dissimilar variables in order to have an equitable comparison. Utility theory is often used for trade studies or analysis of alternatives (AoA), in which multiple attributes must be comparable to each other, in order to provide the decision maker with an "apples-to-apples" comparison. For the purpose of this paper, the inputs and outputs will be scaled between 0 and 1. Other scales may be used (e.g. 0-10, 1-10, 1-100, etc.), but must remain consistent for all factors considered.

Inputs into the utility function will represent the level of uncertainty that will be considered. Our scale will use the value of 0 to represent low uncertainty, and 1 as high uncertainty. An example would be program funding uncertainty, where 0 would represent the program being completely funded, and 1 would represent unclear knowledge of funding. The output of the utility function will represent the impact of uncertainty to the mission execution of the system. The value of 0 will represent a low performance of the mission, and the value of 1 will represent a high performance of

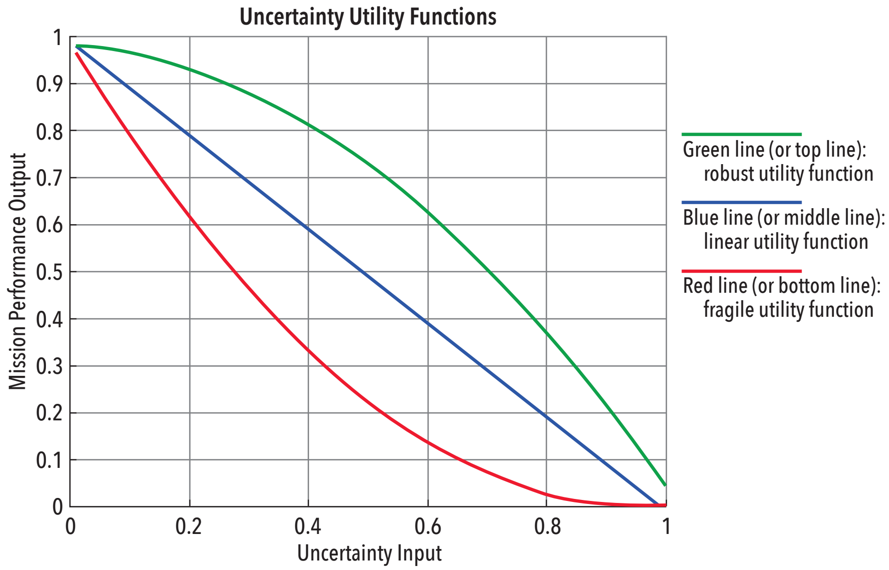


Figure 1. Utility function shape examples

the mission. Continuing our example of program funding uncertainty, a 0 would represent that the effect has no performance in the mission, where a 1 would represent full performance in the mission.

It is envisioned that these utility functions will generally move in a top to bottom, left to right direction to represent the relationship that low uncertainty will contribute to high performance. It may be generated by consulting subject matter experts (SMEs) to determine the shape of these utility functions for each subsystem and the uncertainty impacts on the mission. The relationship may be linear, indicating a gradual decrease of performance as the uncertainty increases. The relationship may be robust, indicating that the performance is not significantly impacted with a slight increase in uncertainty. The relationship may be fragile, indicating that the performance is sensitive to small increases in uncertainty.

Figure 1 provides an example of such a utility function. To determine the shape of these curves, it may be calculated through SME discussion on how sensitive these changes are to the mission performance, or reliance on higher fidelity simulations to develop these utility functions. The simulations could be used in a way to incrementally change one uncertainty level and observe how the other uncertainty levels and mission performance are affected. In this example, it could be a sensor model interacting with a physical environment model and determining the resultant output to mission detection or identification. These uncertainty increments may also represent different maturity levels of the system developments under study. The output of this step is an uncertainty utility function for each unique combination of subsystem/component and uncertainty factors identified in step 1.

Step 3: Describe the Uncertainty Interdependency

This step describes the interdependencies of the uncertainty utility functions, and how one utility function may influence another. This step seeks to identify how some uncertainties contribute to other uncertainties. Some of these dependencies may be one way or two way in nature. An example of such a dependency are multiple subsystems that must perform together in order to complete the mission. In an air-to-air mission, an airborne system must detect the threat (with sensor), identify the threat (with identification (ID) processor), conduct decision making (with battle management (BM) processor) engage the threat (with weapon), and assess the follow-on actions (using communication subsystems). These five subsystems must be able to perform their mission functions, but also in a particular order, and not every subsystem will interface with all other subsystems. Examples of this sequence and interaction require the sensor must be able to complete its detection function in order to then hand off the results to the processor. Similarly, the processor must complete its identification function in order to indicate that a weapon must be launched against a hostile contact. In this example, the processor input has a dependency on the sensor output, and the processor output will then become an input to the weapon. The output of this step is the identification of linkages, to include directionality, between uncertainties.

Step 4: Collect the Uncertainty Inputs (Scenarios)

This step collects the different uncertainties that would affect the system, which may be categorized into scenarios or use cases. These scenarios may represent near-term, mid-term, and far-term threat or environment projections, and may be used to quantify the difference in mission performance based on the changes in uncertainties. Scenarios may be isolated to analyze just the threat, environment, or system uncertainties, or the scenarios may combine multiple uncertainties. The output of this step is the development of multiple scenarios that will be used for analysis.

Step 5: Conduct Overall Mission Uncertainty Analysis

This step will conduct the mission analysis based on the scenario inputs. These scenarios will contain the initial uncertainty levels, which are then injected into the utility functions (shown in Table 2), produce an output to mission performance, and then are linked to other dependent subsystems (shown in Table 3). The process using the interdependency tables and interactions are exercised to represent an elapsed time duration in order to calculate subsystem uncertainty levels. Mission metrics can be calculated through manipulation of the interdependency tables to determine the spread of subsystem uncertainty dependent on the factors (red, weather, and blue). Performing this analysis with different scenarios can show the relative uncertainty performance difference. For the purpose of this paper, each of the uncertainty factors are weighted equally, but in future work, the decision makers may elect to assign different weights to individual factors.

ILLUSTRATIVE EXAMPLE

The five-step methodology is explored with an illustrative example. The example seeks to develop an airborne platform capability that will attempt to detect, identify, prosecute, and engage threat airborne targets, such as aircraft, cruise missiles, or unmanned aerial systems (UAS).

Step 1: Identify the Uncertainty Areas

The example is divided into three uncertainty types: what the threat (red) can do, the operational environment, and the friendly (blue) forces structure and tactics. The example requires five phases of mission execution: the system must search and detect the threat, identify the threat's intentions, decide what actions to take, engage the threat, and assess the next step. These phases correspond to a physical component that will be used to execute that phase. Table 1 provides a top-level view of the uncertainty inputs with a high and low level

Table 1. Example uncertainty levels

Phase	Component	Uncertainty type	Uncertainty input	Inputs (low uncertainty)	Inputs (high uncertainty)
Search/ Detection	Sensor	Threat	Target signature	Good intel	Bad intel
Identify	Identification	Threat	Target jamming	Good intel	Bad intel
Decision Making	Decision Making	Threat	Target Low Probability of Intercept (LPI) communications	Good intel	Bad intel
Engage	Weapon	Threat	Threat weapons	Good intel	Bad intel
Assess	Communications	Threat	Threat tactics	Good intel	Bad intel
Phase	Component	Uncertainty type	Uncertainty input	Inputs (low uncertainty)	Inputs (high uncertainty)
Search/ Detection	Sensor	Environment	Weather conditions to affect detection	Operating in known conditions	Unexpected conditions
Identify	Identification	Environment	Operational conditions to affect identification	Operating in known conditions	Unexpected conditions
Decision Making	Decision Making	Environment	Operational conditions to affect decision making	Operating in known conditions	Unexpected conditions
Engage	Weapon	Environment	Operational conditions to affect engagement	Operating in known conditions	Unexpected conditions
Assess	Communications	Environment	Weather conditions to affect communications	Operating in known conditions	Unexpected conditions
Phase	Component	Uncertainty type	Uncertainty input	Inputs (low uncertainty)	Inputs (high uncertainty)
Search/ Detection	Sensor	Blue	Acquisition changes	Fully funded	Less funded
Identify	Identification	Blue	Acquisition changes	Fully funded	Less funded
Decision Making	Decision Making	Blue	Tactics changes	Operating with known TTP	Operating with different TTP
Engage	Weapon	Blue	Acquisition changes	Fully funded	Less funded
Assess	Communications	Blue	Interoperability changes	Operating with known interoperability	Operating with unknown interoperability

of uncertainty associated with each mission phase and system components.

Step 2: Develop the Uncertainty Utility Function

For this example, we will use a series of three general uncertainty utility functions as described in the methodology, found in Figure 1. The components and mission phases are used from Table 1.

Step 3: Describe the Uncertainty Interdependency

For the purpose of this example, we develop a notional interdependency table between uncertainty factors, provided in Table 2. The initial mapping of our blue system capabilities (y-axis) to the three groupings of uncertainty factors (x-axis) are shown. At each intersection, there are a total of four possibilities: there is no interaction, or one of the three utility function types (robust, linear, and fragile) exists. The compilation of this table would represent the total uncertainty possibilities that the system would encounter. The selection of the uncertainty utility function is dependent on the assessed sensitivity and expected performance of the blue capability against the range of uncertainty factors. This may also be generated through SME assessment or iterative simulation of system performance under a variety of uncertainty factors.

Table 3 provides a notional view of the initial uncertainty factors and their effect on other interdependent uncertainty factors for the air-to-air mission. It is intended to start in the left column, and then read across from left to right to find the contributing inputs. A 0 indicates no contribution/ impact to the mission, and a 1 indicates there is a contributing input to the uncertainty factor. This is performed for both the red (threat) and blue (friendly) subsystems that may interact with each other and thus have a contributing uncertainty. An example is that the variability in target signature

Table 2. Uncertainty interdependency table

	Uncertainty Dependencies									
Blue Capabilities	Target signature	Target jamming	Target LPI comms	Threat weapons	Threat tactics	Weather impacts	Operating environment constraints	Blue acquisition	Blue TTP	Blue interoperability
Sensor	1	2	0	1	1	2	2	2	2	0
Identification	2	1	2	3	3	2	2	2	3	3
Decision Making	2	3	0	2	3	0	2	1	2	1
Weapon	1	1	0	2	1	2	2	1	3	3
Communications	0	1	0	0	0	1	2	2	3	3

Key
0: Not applicable
1: Robust utility function
2: Linear utility function
3: Fragile utility function

Table 3. Uncertainty interdependency factors

	Contributing Uncertainty (read across)				
Initial uncertainty (read down)	Target signature	Target jamming	Target LPI comms	Threat weapons	Threat tactics
Target signature	0	0	0	0	1
Target jamming	0	0	1	1	1
Target LPI comms	0	0	0	0	1
Threat weapons	1	1	0	0	1
Threat tactics	0	0	1	1	0

	Contributing Uncertainty (read across)		
Initial uncertainty (read down)	Blue acquisition	Blue TTP	Blue interoperability
Blue acquisition	0	0	1
Blue TTP	0	0	1
Blue interoperability	1	1	0

Key
0: no contribution
1: contribution

uncertainty will influence red tactics that are employed.

Step 4: Collect the Uncertainty Inputs (Scenarios)

For this example, we will use three different scenarios to evaluate our methodology. The first scenario has a generally good understanding of the threat and an accurate estimation of the uncertainty growth over time, which will be relatively small. The uncertainty utility functions for the threat, environment, and friendly factors will range from 0 (not applicable) to 1 (robust).

The second scenario has an average understanding of the threat, but with a less accurate estimation of the uncertainty. The uncertainty utility functions for the threat, environment, and friendly factors can range from 0, 1, or 2 (linear).

The third scenario has a poor understanding of the threat, and a low estimation of the uncertainty. The uncertainty utility functions for the threat, environment, and friendly factors can range from 0, 1, 2, or 3 (fragile).

Step 5: Conduct Overall Mission Uncertainty Analysis

We are then able to execute our model using the three different scenarios. Figures 2-4 show a boxplot summary of the five system characteristics: sensing, identification, battle management, engagement, and communications, which are shown on the y-axis. On the x-axis, the red systems, weather / operational environment, and blue system uncertainty factors are shown. The boxplot shows the mean (red line), the 1st and 3rd quartile (box), and data within the 1.5 inter quartile range (IQR) of the upper and lower quartiles (whiskers) of the model output. Outliers outside the whiskers are labeled as red crosses. The scenarios were run for 500 timesteps. The last column

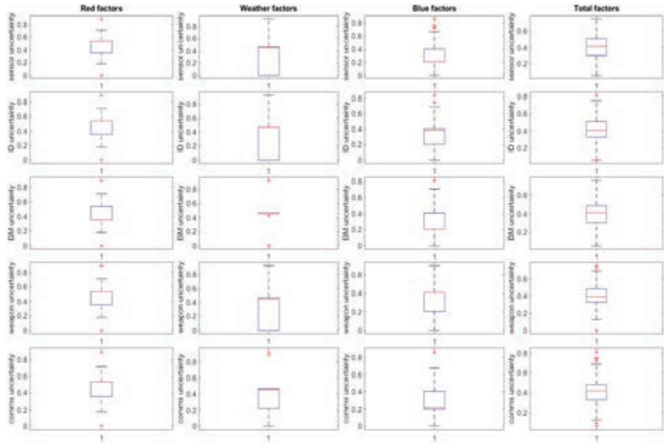


Figure 2. *Scenario 1 analysis*

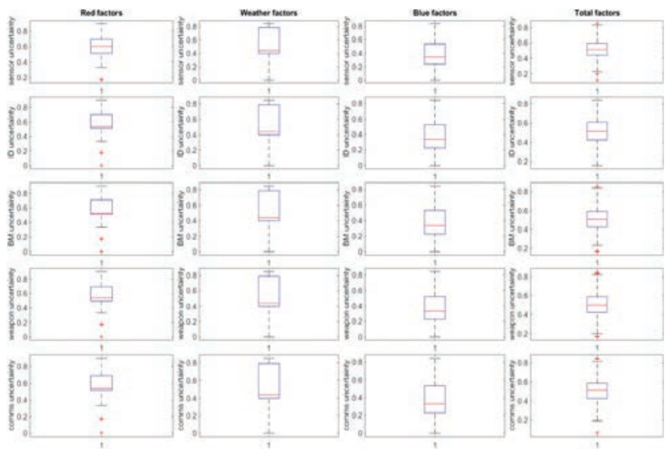


Figure 3. *Scenario 2 analysis*

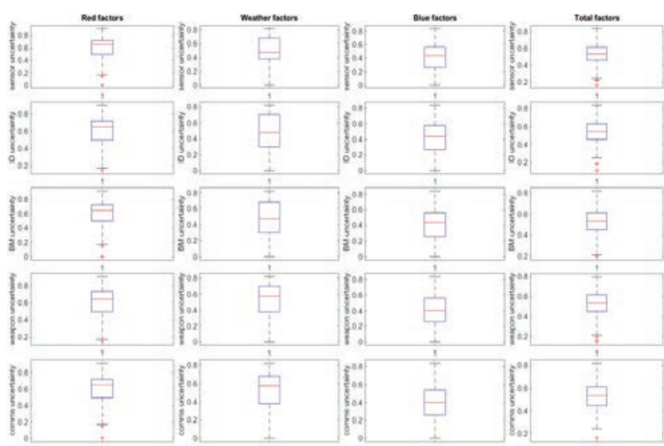


Figure 4. *Scenario 3 analysis*

provides an average of all three factors for the total system performance. Note as the scenarios increase in difficulty, the uncertainty levels increase, as evidenced with the boxplot shapes. This shift towards higher uncertainty levels as more difficult scenarios are introduced would indicate a greater dependency of the mission factors when increased in uncertainty, and are representative of expected trends when faced with more difficult scenarios.

CONCLUSION/FUTURE WORK

This paper has developed a methodology in order to consider uncertainty in terms of three perspectives: uncertainty in the threat performance and employment, uncertainty in the operational environment, and uncertainty in the friendly system interoperability and acquisition. Through use of adjacency matrices and utility functions, we can calculate the relationships between the uncertainty factors and view their interdependent effect on each other as their uncertainty levels change. Decision makers may utilize this approach in order to visualize the difference in uncertainty for system concepts when considering decisions regarding the system development. This level of uncertainty may also correspond to a relative risk comparison of concepts during system development. Future work is reserved for quantifying the level of risk to uncertainty within the system development lifecycle decision making. Additional work would be to apply this methodology to a mission that has available programmatic decision data and compare the validity of the model predictions to the actual outcomes, in order to determine its utility to decision makers.

Future work may evaluate additional programs that have less quantifiable system performance measures (such as emergency management or cyber operations), or a system of systems configuration that may require dependencies between multiple systems in order to accomplish the mission. Having a greater number of systems and interfaces may imply there is a pareto front on the level of expected uncertainty impacts to the overall mission performance. ■

REFERENCES

- Averyt, K., L. D. Brekke, D. E. Busch, L. Kaatz, L. Welling, E. H. Hartge, and T. Iseman. 2013. “Moving forward with imperfect information.” *Assessment of climate change in the Southwest United States: A report prepared for the national Climate assessment*, 436-461.
- Boehm, B. 2009. “Future challenges for systems and software cost estimation.” In Proceedings of the 42nd Annual DoD Cost Analysis Symposium, Feburary.
- Buede, D. M. 2009. The Engineering Design of Systems: Models and Methods. 2nd edition. Wiley.
- Flage, R., and T. Aven. 2009. “On treatment of uncertainty in system planning.” Reliability Engineering & System Safety 94, no. 4: 884-890.
- Kossiakoff, A., W. N. Sweet, S. J. Seymour, and S. M. Biemer. 2011. *Systems engineering principles and practice*. Vol. 83. John Wiley & Sons.
- McManus, H., and D. Hastings. 2004. “A Framework for Understanding Uncertainty and its Mitigation and Exploitation in Complex Systems.” Engineering Systems Symposium, Massachusetts Institute of Technology, Cambridge, US-MA, March.
- Yang, D., B. Boehm, Y. Yang, Q. Wang, and M. Li. 2007. “Coping with the cone of uncertainty: an empirical study of the SAIV process model.” In *International Conference on Software Process*, pp. 37-48. Berlin, Heidelberg, DE: Springer Berlin Heidelberg.

ABOUT THE AUTHORS

David Flanigan is a member of the Principal Professional Staff for The Johns Hopkins University Applied Physics Laboratory, providing systems engineering services to various Department of Defense customers. He currently has over 20 years active and reserve service with the US Navy. A graduate of the University of Arizona, he holds a MS in information systems and technology, a MS in systems engineering from the Johns Hopkins University, and a PhD in systems engineering and operations research from George Mason University.

> continued on page 38

The ValXplore method: exploring desirability, feasibility and viability of business and system design under uncertainty

Sonia Ben Hamida, Marija Jankovic, Alain Huet, and **Jean-Claude Bocquet**

[Editor: This paper refers to the *Systems Engineering Vision 2025* (Copyright 2014 by the International Council on Systems Engineering), INCOSE Systems Engineering Handbook v4 (Copyright 2015 by INCOSE), and ISO 15288:2015.]

ABSTRACT

In early design stages, business developers and systems engineers deal with uncertainties on the business problem, in line with the company's strategy. Before designing the system, the business developers need to set the boundaries of the business problem: What are the values to deliver to which stakeholders? What are their preferences? What are the future trends or the evolution of the markets and the external context? These questions regarding the uncertainties on the definition of the problem may not have answers and need to be investigated to assess the value robustness of the possible design alternatives. The aim of this work is to support decision-making in business and system design thanks to a broad and rapid analysis of a large amount of business design alternatives under uncertainty. We introduce a decision-making support method, called ValXplore, based on visual analysis and data analytics to explore the uncertainties on and in the business problem. The method was tested and validated on an industrial case study to assess the benefits and limits of the semi-reusability of a launch vehicle. Both business developers and systems engineers can rapidly explore a broad space of alternatives to increase the value to the stakeholders, by performing sensitivity and uncertainty analyses.

INTRODUCTION

This research focuses on the concept stage, where business models are built, committed costs are still low, but stakeholders' expectations are often unclear and fuzzy. Decisions in early stages impact between 75% and 80 % of overall system life cost (DAU 2013). Moreover, increase in system complexity is enhancing the need for a more interdependent decision-making process across design disciplines and processes (French 1993, Heisig et al. 2009, Keeney and Keeney 2009, and Roy 2013). The INCOSE systems engineering vision (INCOSE 2014) asks for effective decision making by rapidly exploring a broad space of alternatives to maximize the overall value. Early design stages of complex systems consist in defining the problem space and characterizing the solution space, that is, investigate different concepts regarding multiple objectives like performance, costs, etc.

However, the definition of the business design is often dissociated from the system design. Business and engineering teams work both on eliciting the added values for the customers, but the processes remain separated. Moreover, when developing space systems, stakeholder objectives are often ill formulated or fuzzy. And system architecting becomes difficult to orient as the boundaries of the system are not yet fixed.

That is why defining a common process for business and system design decision-making is essential to gain insight on the best value positioning (desirability) and the technical feasibility of the possible solutions as well as the economic viability of the solution. This triptych needs to be explored conjointly by the business developers, who capture the customers' values and preferences, hand in hand with the systems engineers, who generate solutions and evaluate their performances. This paper introduces the ValXplore method to explore desirability, feasibility and profitability of value propositions under uncertainty, and provide recommendations to the decision makers. The method helps to consider the exogenous uncertainties inherent of a business problem. The proposed method supports the formulation of the business problem, the understanding of the impact of uncertainties on the system architecture, the identification of most valuable system architectures by using trade space exploration. The proposed method was applied to an industrial study at Airbus Safran Launchers on the benefits and limits of the semi-reusable launch vehicle. The method allows the decision makers and engineers to visualize synthesis of the value proposition and the feasible design alternatives, gain insights on the impact of the exogenous uncertainties, and support the formulation of recommendations on the design of both the business problem and the solutions, such as the change of the scope of the value proposition or the update of system architectures.

BACKGROUND

Business problem decision-making support. Traditional systems engineering freezes rapidly the specification of the system of interest and hampers the exploration of the situations to address. The lack of analysis on the uncertainty can dramatically impact the success of the system of interest. More and more the importance of early designs stages is underlined. INCOSE is broadening the scope of systems engineering to address not only engineering activities but also business ones. In the version 4 of the INCOSE handbook (INCOSE 2015), INCOSE added a new process "Business or Mission Analysis" in the concept stage, prior to the stakeholders needs definition. This new process includes the definition of the problem space, but this activity remains little supported today.

The shift from decision theory to decision support methodology highlights the increasing interest in the decision support process (Tsoukiàs 2008). Multi-criteria decision methods (MCDM) focus on the exploration and the evaluation of the alternatives, but not the formulation of the problem (Belton and Stewart 2002). Moreover, the meaning of the weights, that is, the importance of the criterion, is model-dependent (Belton and Stewart 2002). It may be very difficult to elicit the weights in multicriteria decision models.

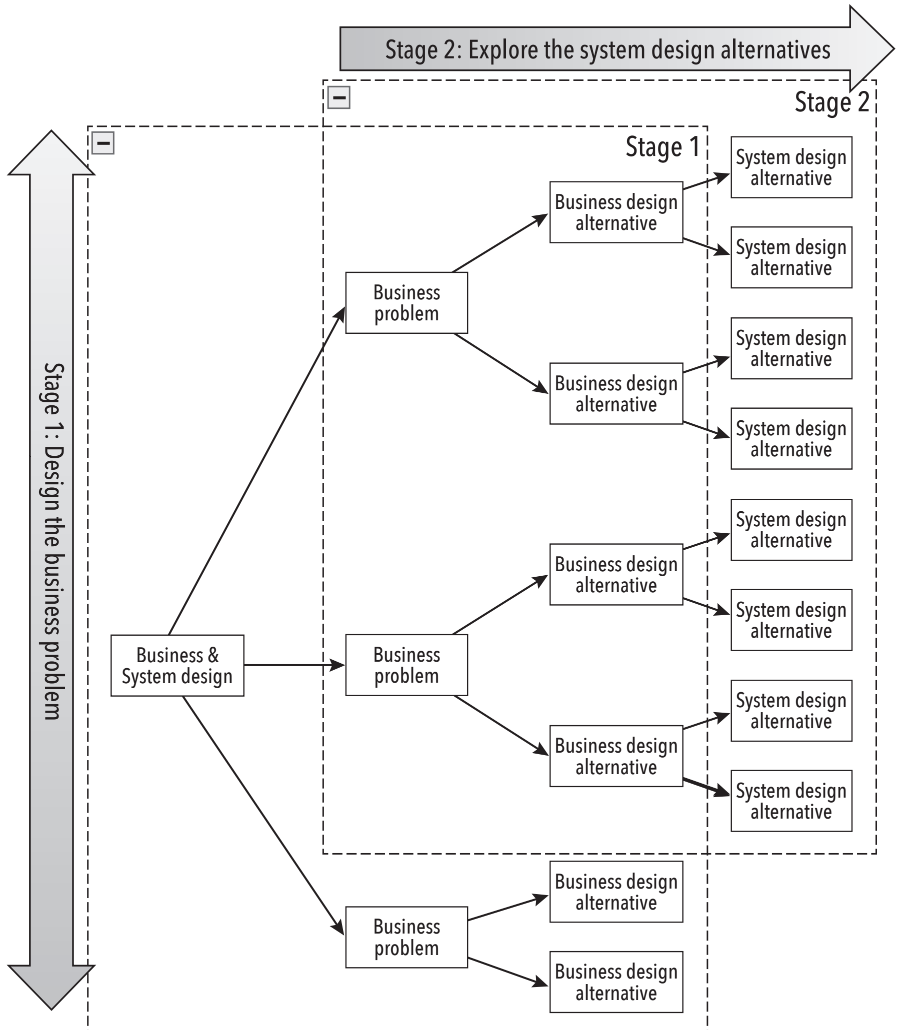


Figure 1. ValXplore method

Two approaches exist to support multiple criteria decisions: (1) Creating a multi-attribute utility function. The function aggregates the different criteria in a single criterion. But the aggregated function may be difficult to interpret by the decision maker. (2) Using pairwise comparison of the alternatives. But the ranking results may be difficult to justify. The two approaches differ in the method to elicit preferences from the decision makers, and the translation of these preferences into quantitative measures.

Decision theory does not consider the real context of the decision (Tsoukiàs 2008): who decides, who are the stakeholders, what is the quality of the information, the level of uncertainty, etc. Moreover, the impact of the decision support process on the decision appears to be more important than the applied method itself (French 1993, Keeney and Keeney 2009, and Roy 2013). For Simon (1983)

"a decision is not an act, but a process". The decision process can become complex when the decision problem involves several stakeholders carrying different values and preferences.

Roy (2013) explains what is missing in MCDM and what is expected from a decision support methodology: Determining how to formulate a problem, determining the preferences of the decision makers, aggregating multiple criteria preferences, and developing recommendations.

System architecture trade space exploration. The term trade space is a combination of the words

"trade-off" and "play space". A trade space is an "area of evaluation bounded by a prescribed set of boundary constraints that serve to scope the set of candidate alternatives for further trade study analysis" (Wasson 2015). The trade space exploration is described as a shopping process where the decision makers discover what they want while they are looking for it. Ross et al. (Rader et al.) investigate the value robustness of a system. Ross and Rhodes (2008) define value robustness as "the ability of a system to continue to deliver stakeholder value in the face of changing contexts and needs". The value of the system is examined with regard to possible future contexts. For example, the operating environment of the system, the stakeholders' preference, the market demand, the competitive forces, the technologies' maturity, or the regulatory environment can evolve throughout the lifecycle of the system. The system's value robustness is assessed with regard to these exogenous uncertainties.

Decision-making uncertainties management. Designing complex systems requires to understand the possible contexts where the system will operate (Rhodes and Ross 2010). The economic conditions, policies, and markets may evolve. These exogenous uncertainties need to be explored because they will drive the business design decisions. Traditional systems engineering describes the system boundaries, external systems, external interactions, and the concept of operations but do not support a prescriptive analysis to support decision-making on the orientation of the business design. French (1995) identifies 10 different sources of uncertainty in the decision problem formulation. He groups them into problem structuring, exogenous uncertainties exploration, and results interpretation. Browning et al. (2006) explain the implications of uncertainties on the development of complex systems. The product development activities will vary depending on the level of uncertainty.

RESEARCH DESIGN

We undertook our study within Airbus Safran Launchers. We applied the design research methodology (Blessing and Chakrabarti 2009) to develop and validate the ValXplore method.

Research clarification. The research focuses on business design in early design stages. We undertook a comprehensive study of the existing situation by conducting a series on interviews of two business developers and six system engineers at Airbus Safran Launchers in 2014 (Summers and Eckert 2013). The main questions and hypotheses were defined based on both the interviews and the documentation analysis of in-house processes and projects deliverables.

Descriptive study I: understand design. We observed the concurrent engineering sessions, recorded team discussions and activities of the project detailed in the case study section. Sixteen (16) concurrent engineering sessions were performed, involving 15 disciplines. The documents of the project were analyzed to understand the activities and deliverables realized by the team.

Prescriptive study: develop design support. The ValXplore method was applied to industrial projects within Airbus Safran Launchers. Each step and output of the method was recorded. A review of the research tools in visualization was done based on the most influential research in visualization identified by the IEEE VIZ community and summarized the VIS25 timeline (Rhyne et al. 2015). The commercial solutions were also assessed based on Gartner's magic quadrant on business intelligence (Gartner 2016a) and advanced analytics platforms (Gartner 2016b). The relevant tools to support the method were compared and selected. The project's post-mortem review was organized to identify the main successful and unsuccessful elements about the methods, the tools, and the organization.

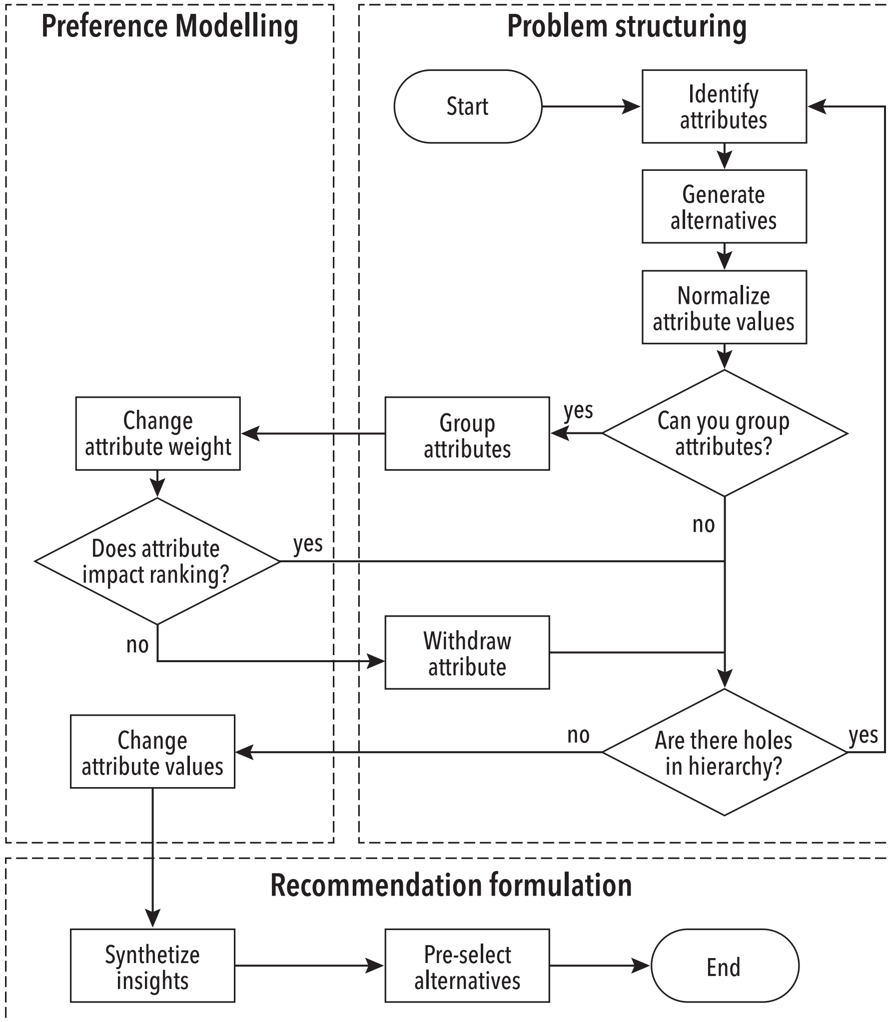


Figure 2. Stage 1: Design business problem – flowchart

THE VALXPLORE METHOD

The ValXplore is a two-stage decision support method to structure and explore of the business design problem and the relevant system design solutions, see Figure 1. The decision maker will learn and understand what is possible (feasibility), what is preferred (desirability) and what matters (viability).

Stage 1: Design Business Problem

The goal of the stage 1, described in Figure 2, is to develop a common understanding of the business design problem, that is, to define the decision to make and the criteria to evaluate the alternatives. We assume the decision makers do not have a clear idea of the problem (Moscarola 1984). The questions supported are: What are the objectives and attributes? What are the preferences of the decision makers? We propose to do a sensitivity analysis on the design of the business problem. This activity requires to generate a rich number of alternatives. A shortlist of potential alternatives will be selected at the end of the first stage to investigate in more details their strengths and weaknesses.

For this first stage, we use the research tool LineUp (Gratzl et al. 2013 and Gratzl 2014) to create, visualize, and explore ranking of the business design alternatives, and perform a visual analysis of the multi-criteria decision problem. The visual analysis helps to interpret the ranking, rapidly compare and analyse alternatives rankings, and

Table 1. Variables of the design problem explored in stage 1

Term	Notation	Definition and equation
Alternatives	$\vec{x}$	$\vec{x} = \begin{pmatrix} x_1 \\ x_i \\ x_m \end{pmatrix}$
Attribute values	A	$A = (a_{ij})_{1\le i\le m, 1\le j\le n}$
Filter range	$[f_{j_{min}}, f_{j_{max}}]$	Filters $f_{j_{min}}$ and $f_{j_{max}}$ on the attribute vector $\vec{a_j}$ to remove alternatives with attribute value a_{ij} outside the filter range $[f_{j_{min}}, f_{j_{max}}]$ from the ranking.
Mapping function	m_j	$m_j : a_{ij} \to [m_{j_{min}}, m_{j_{max}}] \mid 0 \le m_{j_{min}} \le m_{j_{max}} \le 1$
Mapped attribute values	A'	$A' = m(A) = (a'_{ij})_{1\le i\le m, 1\le j\le n}$
Hierarchy level	l	Number of levels in the hierarchy
Hierarchy level weights	W_k	W_k is the weight assigned to the aggregated attributes of level k: $W_k = (w_{ij})_{1\le i\le m, 1\le j\le g_{k-1}} \mid 0 \le w_{ij} \le 1 \wedge \sum w_{ij} = 1$ Where g_{k-1} is the number of groups at level $(k-1)$
Alternative score	s	$\vec{s}(\vec{x}) = A' \prod_{k=0}^{l-1} W_{l-k}$

Table 2. Example of attribute values

Alternative	a01	a02	a03	a04	a05	a06	a07	a08
x01	0.2		TRUE	high	risus	1.34 €	1	21.4
x02	0.2	2.5	TRUE	high		5.53 €	0	78.7
x03	0.7	7.6	FALSE	medium		5.67 €	1	99.1
x04	3.4	2	FALSE		et	5.69 €		45
x05		0.6	TRUE	low	nisi	9.35 €	1	11.2
x06	9.8	8.8	FALSE	high	at	8.59 €	1	

understand how the multiple heterogeneous attributes affect the ranking. The decision makers can interactively combine attributes and refine parameters to explore the effect of changes in the attribute combination, and gain insights on the problem formulation. The stage 1 encompasses value-focused and alternative-focused thinking. The outcomes are the formulation of the business design problem, and a shortlist of alternatives further explored in stage 2.

Methods using hierarchies usually propose a top-down approach to define the objectives and refine them, then model the preferences and evaluate the alternatives. In the first stage, we do these three steps all-in-one to give more insights to the decision makers on the problem formulation. Table 1 lists all the variables characterizing the design problem and that we explore.

Problem structuring.

Identify attributes. An objective indicates the direction the decision makers would like to go while an attribute α measures the achievement of the objective. For example, the objective "minimize time-to-orbit" is measured with the attribute "days". The attribute gives the information to understand and assess if the associated objective is achieved. In this step, the decision makers list their objectives and associated attributes. They express what they want, value, and their constraints. Bond et al. (2010) identify two obstacles to generate objectives: "not thinking broadly enough about the range of relevant objectives, and not thinking deeply enough to articulate every objective," and recommend to use a list of possible objectives to identify additional relevant objectives. In the next steps, we propose to explore different combinations of attributes to overcome these issues.

Generate alternatives. The designers generate a rich number of potential alternatives $\vec{x}$. They ask themselves, for example: What could be the perfect, terrible, and reasonable alternatives? They evaluate the attribute values with quantitative performance models or with subjective expert judgements. Missing values can be inferred by computing mean and median or with more complex algorithms with the tool LineUp. The decision makers are visually aware of the missing data with a dashed border inside the bars. Note that the list of alternatives will evolve with the design of the business problem, as this is a search and learning process increasing awareness on the objectives of the design problem. See Table 2.

Normalize attribute values. The objective is to compare the attributes $\vec{a_j}$ with each other. After importing the data into LineUp, the attribute values are normalized, that is, the attribute values are mapped to the interval [0,1], where 0 is "of no interest" and 1 "of interest". It is possible to test different normalizations by changing the mapping function m, and instantly see the effect on the ranking. The decision makers can analyse the distribution of the attribute values, to understand to what extent the attribute discriminates the highest values.

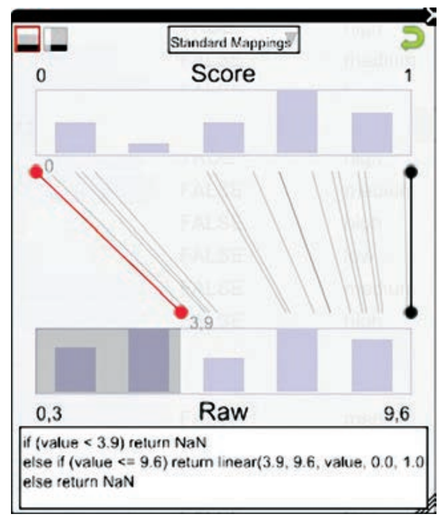


Figure 3. Normalize attribute values – mapping function and filter range

To exclude alternatives $\vec{x}$ not compliant with constraints, filter ranges $[f_{min}, f_{max}]$ can be applied on the attributes A. For example, the decision makers may want to exclude alternatives not compliant with regulations.

Can you group attributes (lateral extension of the hierarchy)? The decision makers combine attributes to construct a weighted sum and sort the alternatives.

Group attributes. The decision makers try out different hierarchies to structure the list of identified attributes in a meaningful way and gain further insight on the problem by comparing the alternatives' rankings with different attribute combinations. See Figure 4.

Are there holes in the hierarchy? The designers can ask themselves what is good or bad about each alternative. Are the

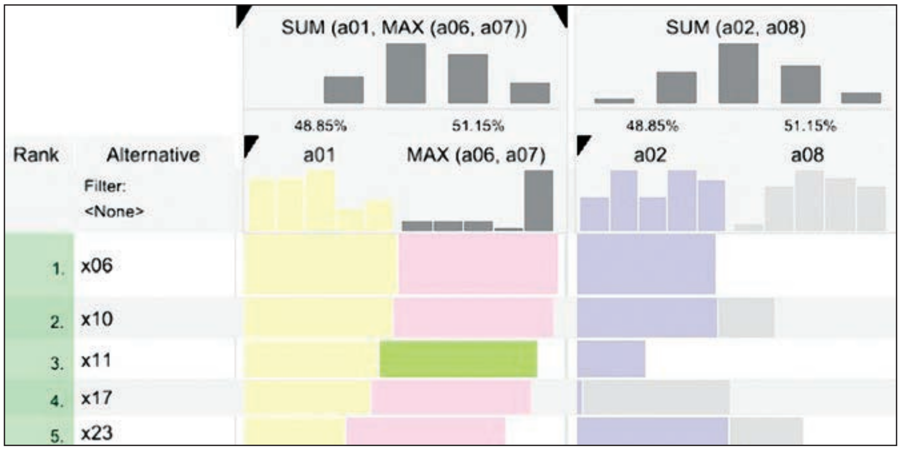


Figure 4. Group attributes

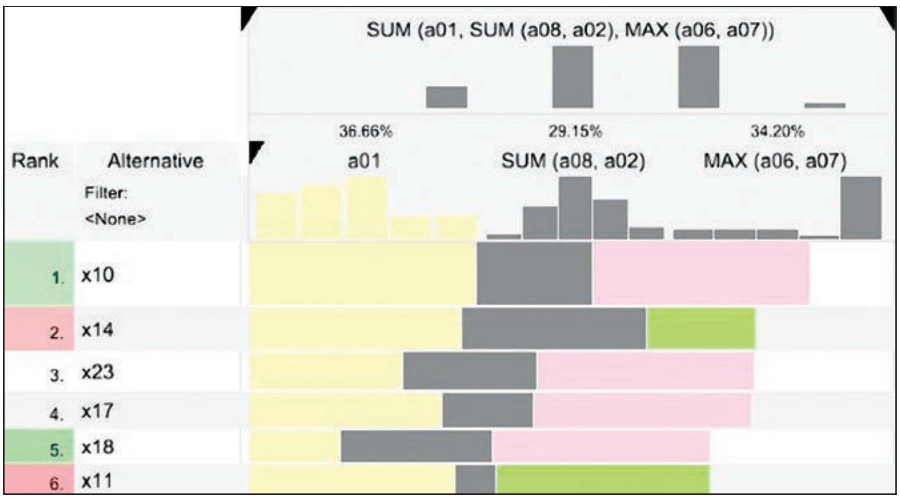


Figure 5. Change attributes weight, visual impact on ranking

strengths and weaknesses of the alternatives captured through the identified attributes? If not, identify the missing attributes and add them.

An attribute can be added in several weighted sums by duplicating the attribute column. For example, the attribute "price" is relevant for all the customer segments, with possibly different preferences.

Preference modelling. The preferences of the decision makers can be captured through many ways, such as market research, focus groups or interviews with the stakeholders about possible contexts of use. However, conflicting preferences may exist making hard to aggregate preferences and maximize value, and preferences may be fuzzy for unarticulated needs. In this step, we consider individual stakeholder preferences and how they may vary across stakeholders. We explore changes in stakeholders' preferences that can occur in response to context shifts, like economic changes, market growth evolutions, threats, etc. French (1995) identifies two types of uncertainties related to preference modelling: (1) Uncertainty about the evolution of future beliefs and preference: For example, what are the possible evolution of the stakeholders' preferences? (2) And uncertainty about judgements. We propose to explore both uncertainties by interactively combining criteria and interpreting the effect of these changes in the criteria combination.

Change attributes weight. The preferences are defined by weights associated to hierarchy level weights, W_k, which group both attributes' weights and weighted sum weights. This step consists in changing the weight of one or more attributes to understand how the attributes influence the ranking of the alternatives. See Figure 5. The decision makers can explore stakeholders' preference changes over time or stakeholders' relative importance regarding the company's strategy to simulate, for example, market growth evolution.

Does the attribute impact the ranking? The decision makers change attributes weight and check if the ranking is impacted. Guiding questions:

- How far to decompose the attributes (vertical extension)? The weights of the lowest attributes (leaves) of the hierarchy are changed to see if it impacts alternatives' ranking.
- For each attribute, does the selection of the alternative could be altered if the attribute was excluded? If not, withdraw the attribute.

Change attribute values. The attribute values may involve uncertainties and judgmental imprecisions. In this step, we propose to adopt an alternative-focused thinking to look at the strengths and weaknesses of the relevant alternatives. The decision makers can explore the effect of changes in attribute values or optimize the values and weights to find the best possible ranking of a specific alternative. See Figure 6.

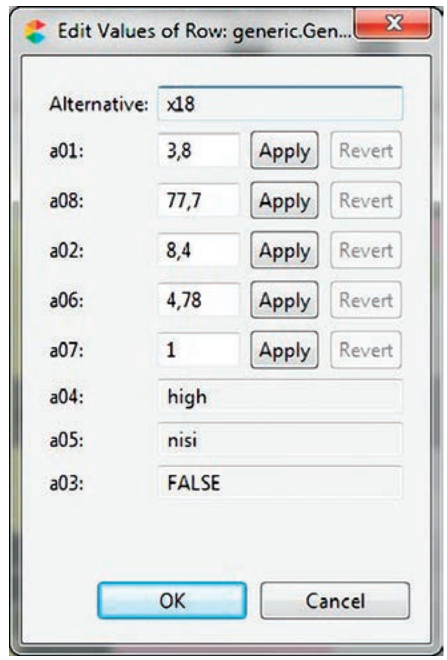


Figure 6. Change attribute values

Recommendation formulation

Synthesize insights. Each step helps the decision makers understand and explore their beliefs, perceptions, and preferences and form and evolve their judgments. LineUp affords to take snapshots of the settings. We suggest saving the meaningful settings that help the decision makers to gain insights on the robustness of the top ranked alternatives over a range of possible futures. For example, what is the robustness of the final ranking? What attributes combination and weighting give the same ranking and affects the final ranking? What attribute values highly impact the ranking? These values may require a more in depth evaluation of the alternatives' attribute values. See Figure 7.

Pre-select alternatives. Select a shortlist of the top-ranked alternatives.

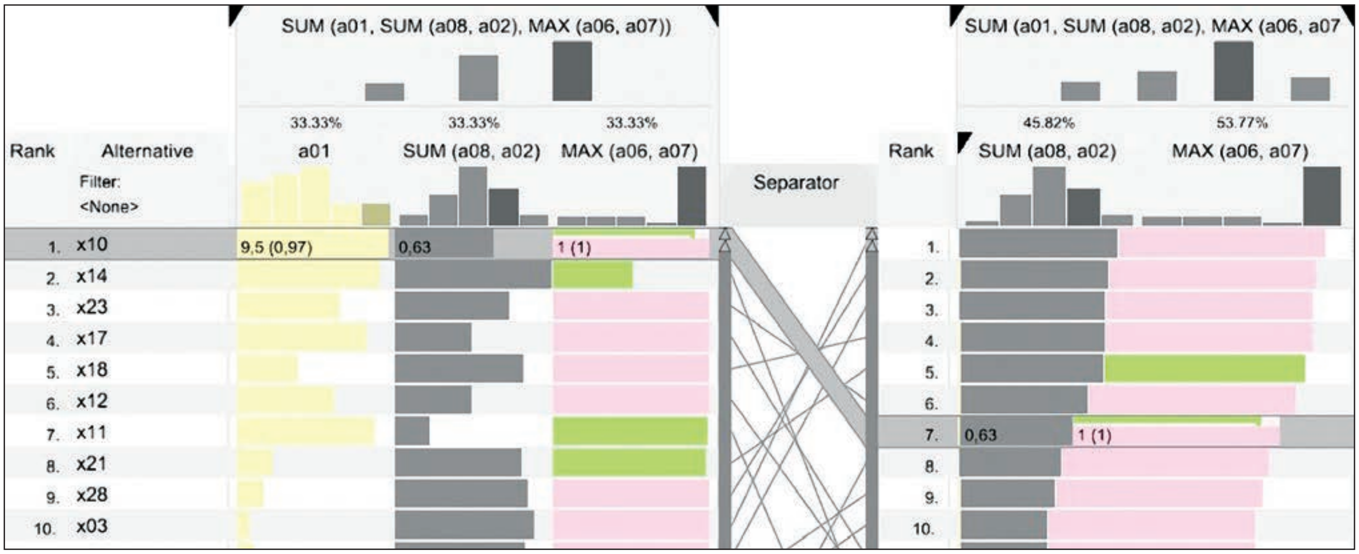


Figure 7. Synthesize insights – settings' snapshots

Stage 2: Explore Business and System Design Alternatives

In stage 1, we selected a shortlist of alternatives for more detailed investigation and evaluation in stage 2. See Figure 8.

Define the possible futures. In this step, the exogenous uncertainties are characterized. A scenario is a what-if story used to explore critical future uncertainties. Scenarios do not aim to predict the future and are based on knowledge from the past and the present. They help to examine the plausible futures – such as the worst, the most likely, and the best cases – to understand the range of possible outcomes. Scenario analysis helps to consider high uncertainties and to identify potential challenges.

We propose to first establish the scope and the focus of the scenarios and to identify the factors and their positive or negative influence. The wider the range of solicited experts, the more exhaustive the identification of scenarios. Then, the most influential factors are identified. For each critical uncertainty, the plausible alternatives and assumptions are identified: What is assumed in this scenario? What assumptions need to be made to arrive to this scenario but are missing? How good are these assumptions? What-if an alternative assumption is made?

Define the business and system design variables. We consider both business and system design variables. Business variables refer, for example, to the value proposition, the customer segments, the price (margin), etc. We propose to apply design structure matrices (DSM) in concurrent engineering (CE). Today, the DSM are applied to the system to increase the pace of work by bringing together the relevant disciplines.

Understand how the business and system design variables are correlated. A scatter plot displays the correlation between a pair of variables, while the regression analysis quantifies the relationship among two or more variables. Scatter plot matrices are constructed to understand the correlation between several variables, identify trade-off, and possible missing variables to characterize the problem and compare the design alternatives.

Identify feasible design alternatives.

Evaluate design alternatives'

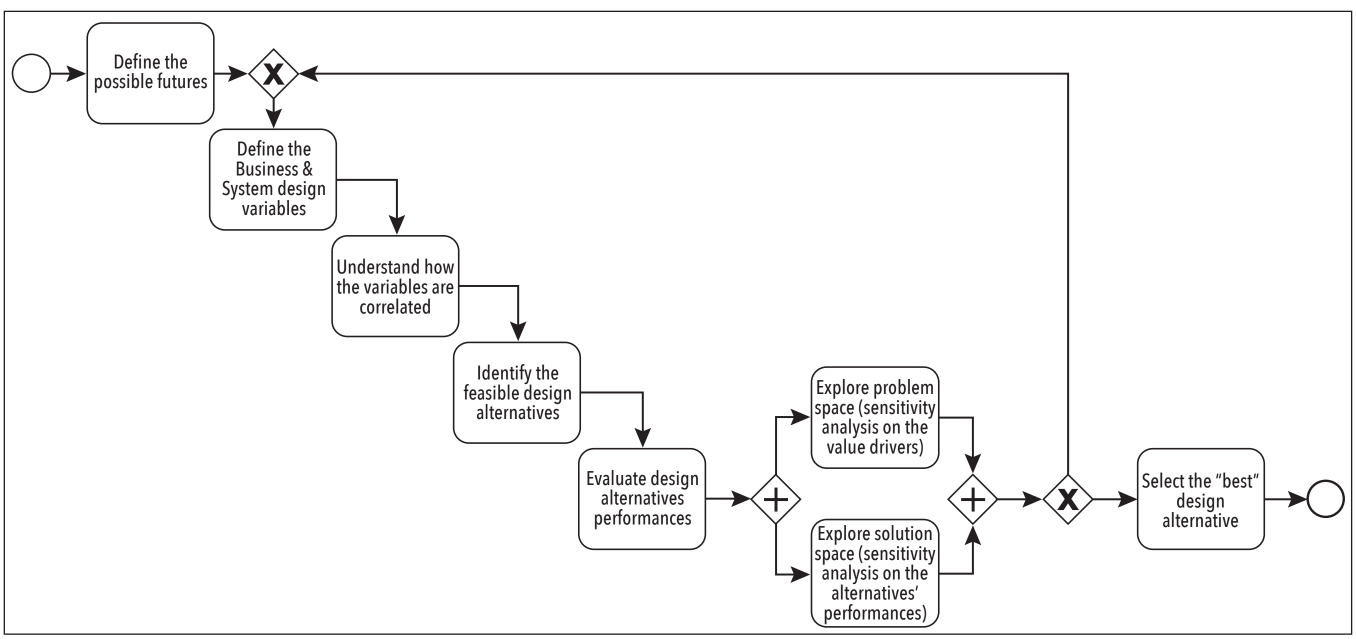


Figure 8. ValXplore stage 2 steps

performances and cost. Develop performance and cost models to evaluate the performances of the design alternatives.

Explore solution space (sensitivity analysis on the alternatives' performances). Perform sensitivity analysis on the design variables.

Explore problem space (sensitivity analysis on the value drivers). Change the variables describing the value proposition, the target customers, etc. Change the exogenous uncertainties to understand the impact on the design alternatives.

Provide recommendations. Example of recommendations: change the value proposition, optimize a design alternative. Refine a performance or cost model. We select the best design alternative regarding changing contexts.

INDUSTRIAL CASE STUDY: THE SEMI-REUSABLE LAUNCH VEHICLE

The proposed method was applied on an industrial project at Airbus Safran Launchers. The goal of the project is to understand what the benefits and the limitations of various reuse options for a launch vehicle are. The business design problem consists in understanding the conditions where reusability of the launch vehicle brings value to the future customers and their potential needs, including for example various targeted orbits and payload constitutions. Is it worth it to invest in such or such reusable system?

The project involves the institutional customers. The project team gathers a dozen of experts from the business development, system engineering, re-entry, costing, design office, mission analysis, and propulsion departments at Airbus Safran Launchers.

Stage 1: Design the business problem

For stage 1, we used the demo version of LineUp, freely available at http://lineup.caleydo.org. The decision makers expressed the need to "reinforce the selection of the decision criteria."

Problem structuring.

Identify attributes. Market research was done to identify a list of potential objectives and attributes. Over twenty values were identified with regard to the considered customer segments.

Generate alternatives. The designers did an extended literature review and identified a wide set of reuse concepts. The performances were assessed based on documentation of expert judgement.

Normalize attribute values. The attribute values were mapped to the interval [0,1], where 0 is "of no interest" and 1 "of interest". Some "killing" attributes were identified.

Group attributes. The decision makers discussed the potential adding and withdrawing attributes. They identified key objectives and discussed which attributes could measure their achievements. They tested several hierarchies, i.e., the attributes of interest and the way to group them.

Preference modelling.

Change attributes weight. The decision makers took time to set their preferences as they add divergent objectives. They were unsure about the relative preference of some attributes, and the selection of the alternatives. can be captured through market or through interviews with the stakeholders about possible. Changing in the weighting helped to justify the importance of the attributes. When lowering the importance of one the attributes, some surprising alternatives ranked on top, and it helped to understand that non-economically viable solutions could be wrongly selected if this attribute's weighting was too low.

Change attribute values. Some attribute values raised discussion about their possible imprecisions. The experts assed the alternatives and, to reach consensus, the decision makers explored the effect on the ranking of changes in some attribute values, such as the technical readiness level of the alternatives. The designers also explored how to optimize the values and weights of preferred alternatives to understand their strengths.

Recommendation formulation.

Synthesize insights. Screenshots of the settings were captured with LineUp to capture the robustness of the top ranked alternatives.

Pre-select alternatives. The decision makers selected the shortlist of alternatives for further evaluation. See Figure 9.

Stage 2: Explore the system design alternatives

The objective of this stage is to explore what is possible and what is not. A more in-depth analysis is performed to assess the risks and opportunities of the selected alternatives.

Define the possible futures. Three market scenarios were identified to consider the uncertainties on market demand, such as the launch of big constellations.

Define the business and system design variables. The design structure matrix (DSM) of the launch vehicle was filled in to understand which discipline needs which information. The data flows were defined from the optimized DSM. More than forty system design variables were identified by the engineering team such as the configuration of the vehicle (number of stages and boosters), the booster diameter, the propellant type (liquid, solid), etc. The business variables are for example the pricing strategy, the market coverage, the launch cadence.

The engineering team worked in concurrent engineering sessions every week to set up and evaluate the feasibility of the systems architectures.

Understand how the business and system design variables are correlated. Scatter plot matrices were built up with the data analytics software Tableau©. New variables were identified from this analysis to better depict the relationship between the market scenarios and the systems' performances.

Evaluate design alternatives' performances and cost. Quick loops were designed to rapidly evaluate the feasibility of the design alternatives. Steele et al. (2002) developed examples of SRLV performance models.

Explore problem space. A sensitivity analysis was performed on the cost drivers.

Explore solution space. Different value propositions were studied and the adaptability to market of the fleets.

Provide recommendations. The exploration helped to understand the strengths and weaknesses of the three alternatives selected. The team decided to withdraw on the alternatives and further explore the two remaining ones. The exploration helped to identify recommendations on the following axes:

- Refine parts of the economic model,

Figure 9. Illustration of case study stage 1

- Explore most important cost drivers,
- Improve the performances of the fleet (system architecture changes),
- Understand the conditions where reusability is most and less interesting.

CONCLUSION

Business design and system design are often separated activities in early design stages, although they are interlinked. We propose a method to explore the desirability, feasibility and viability of business and system design under uncertainty. We characterized the uncertainties on the business problem and defined a first stage to explore these uncertainties to gain insights on structuring the problem and to rapidly assess the value robustness of the design alternatives. A shortlist of alternatives is then selected to further refine the design. In stage 2, we propose to extend the boundaries of the design exploration to the business design by using data analytics.

The method was successfully validated on an industrial project and showed how it could support the understanding of the benefits and limits of a business case. The project team acknowledge that "decision criteria cannot be fixed since the beginning because stakeholders, facing options, learn gradually what they in fine expect and prefer". The project team was satisfied by the method to "ease the understanding of each discipline's contribution" and "ease the communication between the business and engineering teams." The decision makers could gain insight on the design problem in a short period of time. However, the analysis is dependent on the quality and reliability of data (Gordon 2008) and concerns were expressed on carefully interpreting the results. ■

REFERENCES

- Belton V, and T. Stewart. 2002. *Multiple criteria decision analysis: an integrated approach*. Springer Science & Business Media.
- BKCASE Editorial Board. 2016. "The Guide to the Systems Engineering Body of Knowledge (SEBoK), v. 1.6". The Trustees of the Stevens Institute of Technology, Hoboken, US-NJ.
- Blessing L. T., and A. Chakrabarti. 2009. DRM: A Design Reseach Methodology. Springer, London.
- Bond S. D., K.A. Carlson, and R. L. Keeney. 2010. "Improving the generation of decision objectives." Decision Analysis 7 (3):238–255.
- Browning T. R., E. Fricke, and H. Negele. 2006. "Key concepts in modeling product development processes." *Systems Engineering* 9:104–128.
- DAU. 2013. *Defense Acquisition Guidebook*. Defense Acquisition University Press.
- French, S. 1993. "A manifesto for the new MCDA era." Journal of Multi-Criteria Decision Analysis 2 (3):125–127.
- French, S. 1995. "Uncertainty and imprecision: Modelling and analysis." *J Oper Res Soc* 46:70–79.
- Gartner. 2016a. Magic Quadrant for Business Intelligence and Analytics Platforms. https://www.gartner.com/doc/reprints?id=1-2XYY9ZR&ct=160204&st=sb. Accessed 17 Mar 2016.
- Gartner. 2016b. Magic Quadrant for Advanced Analytics Platforms. https://www.gartner.com/doc/reprints?id=1-2YC9GD6&ct=160209&st=sb. Accessed 22 Apr 2016.
- Gordon, A. 2008. *Future savvy: identifying trends to make better decisions, manage uncertainty, and profit from change.* AMACOM Div American Mgmt Assn.
- Gratzl, S. 2014. LineUp: Multi-Attribute Rankings. http://www.caleydo.org/tools/lineup/. Accessed 14 Oct 2016.
- Gratzl, S., A. Lex, N. Gehlenborg, H. Pfister, and M. Streit. 2013. Lineup: *Visual analysis of multi-attribute rankings.*" Vis Comput Graph *IEEE transactions on visualization and computer graphics* 19 (12): 2277-2286.
- Heisig, P., J. Clarkson, J. Hemphälä, C. Wadell, M. N. Bergendahl, J. Roelofsen, M. Kreimeyer, and U. Lindemann. 2009. "Challenges and future fields of research for modelling and management of engineering processes." 1-40.
- INCOSE. 2014. *A world in motion-systems engineering vision 2025*, International Council on Systems Engineering.
- INCOSE, ed. 2015. INCOSE *Systems Engineering Handbook*, 4th ed. John Wiley & Sons.
- Keeney R. L. 2009. *Value-focused thinking: A path to creative decisionmaking.* Harvard University Press.
- Moscarola, J. 1984. "Organizational decision processes and ORASA intervention." In *Rethinking the Process of Operational Research & Systems Analysis*, pp. 169-186. Pergamon.
- Rader A., A. M. Ross, and M. E. Fitzgerald. 2014. "Multi-Epoch Analysis of a Satellite Constellation to Identify Value Robust Deployment across Uncertain Futures." In *AIAA SPACE 2014 Conference and Exposition*, p. 4269.
- Rhodes D. H., and A. M. Ross. 2010. "Five aspects of engineering complex systems emerging constructs and methods." In 2010 IEEE International Systems Conference, pp. 190-195. IEEE.
- Rhyne T-M., J. Stasko, H. Hagen, and B. Bach. 2015. VIS 25 Timeline. http://www.aviz.fr/~bbach/vis25timeline/. Accessed 21 Nov 2016
- Ross, A.M., and D. H. Rhodes. 2008. "Architecting systems for value robustness: research motivations and progress." In *2008 2nd Annual IEEE Systems Conference*, pp. 1-8. IEEE.
- Roy, B. 2013. *Multicriteria methodology for decision aiding.* Springer Science & Business Media.
- Simon, H. A. 1983. "Search and reasoning in problem solving." *Artif. Intell.* 21, no. 1-2: 7-29.
- Steele, M. J., M. Mollaghasemi, G. Rabadi, and G. Cates. 2002. "Generic simulation models of reusable launch vehicles." In *Proceedings of the Winter Simulation Conference*, vol. 1, pp. 747-753. IEEE.
- Summers, J. D., and C. Eckert. 2013. Interviewing as a method for data gathering in engineering design research. Clemson Engineering Design Applications and Research.
- Tsoukiàs, A. 2008. "From decision theory to decision aiding methodology." *Eur J Oper Res* 187:138– 161.
- Wasson, C. S. 2015. *System Engineering Analysis, Design, and Development: Concepts, Principles, and Practices*. John Wiley & Sons.

ABOUT THE AUTHORS

[Editor: Author biographies were current when the paper was initially submitted for publication in 2017.]

> continued on page 38

Informing the delineation of input uncertainty space in exploratory modelling using a heuristic approach

Enayat A. Moallemi, Sondoss Elsawah, and **Michael J. Ryan**

ABSTRACT
Exploratory modelling is an emerging approach which can address the challenge of model-based decision making in dealing with input model uncertainties. Exploratory modelling samples from an input uncertainty space and generates extensive computational experiments to analyse possible model behaviours in an output solution space. The way that the input uncertainty space is delineated influences the results of exploratory modelling and its computational cost. In this article, we show the statistical significance of the implication of the size of an input uncertainty space on the resulted output solution space. We also propose a heuristic approach which informs the delineation of input uncertainties by screening the relevant model behaviour in the solution space. An illustrative example of an aircraft fleet management system is used to demonstrate the implementation of our approach in practice. We conclude that the delineation of input uncertainty space can be a way to control simulations in exploratory modelling and to enhance the efficiency of the exploration process and the confidence of the final results.

INTRODUCTION

Models are used widely in the decision science to represent a system of interest and to assist in the decision-making process. The analysis of model behaviour for decision making is challenged by the presence of various forms of uncertainty. Exploratory modelling is an emerging field which addresses this challenge through exploring the implications of various possible values of input uncertainties on the model behaviour (Bankes 1993, Kwakke, 2017, and Moallemi *et al.* 2017). Exploratory modelling uses one or more simulation models to generate the possible impacts of the input uncertainty space on the output solution space in the format of an ensemble of computational experiments (Bankes *et al.* 2001). The design of these experiments—in terms of which uncertainty factor to choose, how to delineate the uncertainty space, how to sample from the uncertainty space, and how many samples to collect—is critical in the exploratory modelling process (Pianosi *et al.* 2016, Schulze *et al.* 1999, and Bergmann *et al.* 2011). Experimental design is critical as different designs can result in different ensembles of generated experiments and varied output solution spaces, and therefore, can sometimes result in divergent decision insights (Pianosi *et al.* 2016, and Kwakkel and Pruyt 2015).

Previous studies have discussed the aspects of the design of experiments in exploratory modelling to different extents (for example Lempert *et al.* 2003, Kwakkel *et al.* 2010, and Haasnoot *et al.* 2013). In this article, we focus on one of these aspects: the delineation of the input uncertainty space. We assume that the behaviour of the output solution space is dependent on the areas of the input uncertainty space from which samples are taken, and therefore, not all areas of the input uncertainty space should be investigated if only a specific behaviour in the output solution space is of interest to decision makers. Accordingly, this article answers the two following questions:

- Question 1: How sensitive is the output solution space to changes in the input uncertainty space?
- Question 2: How can we inform the delineation of the input uncertainty space by screening a behaviour of interest in the output solution space?

We first show the significance of the input uncertainty space on the distribution of the output solution space, with an illustrative exploratory modelling example in the aircraft fleet management system. We then propose and demonstrate a heuristic approach to inform the delineation of the input uncertainty space based on screening output solutions. An informed delineation of input uncertainties can reduce the computational burden of the exploratory modelling process by avoiding extra simulations based on the samples from the irrelevant areas of input uncertainty space. This screening also presents a more detailed picture of the particular area of interest within the solution space.

The remainder of the article is structured as follows. The second and third sections present an overview of the research background and methods used in this work. The fourth section answers the two research questions posed. The final section concludes the paper and draws future research directions.

BACKGROUND

This section gives a broad overview of exploratory modelling and why experimental design, of which the delineation of uncertainty space (the focus of this article) is one part of design, that is critically important.

EXPLORATORY MODELLING

Exploratory modelling (Lempert *et al*. 2003, Bankes *et al*. 2001, and Bankes 1993) challenges the reliance on deterministic (or even probabilistic with known distributions) sets of model structures and input parameters in modelling in the face of future uncertainties. Exploratory modelling supports the exploration of the impacts of a diversity of parametric and non-parametric assumptions on the output solution space, when the model operates in 'deep or severe uncertainty' (Lempert et al. 2003, and BenHaim 2006). Exploratory modelling generates various model responses in thousands of computational experiments, using a simulation model and sampling from an input uncertainty space. It then analyses the generated experiments with a range of analytical techniques to draw various decision insights and modelling conclusions (Moallemi *et al*. 2017, Kwakkel 2017, and Lempert 2013). See (Moallemi *et al*. Submitted-a, and Walker *et al*. 2013) for a further explanation of exploratory modelling. Exploratory modelling shares similarities with sensitivity analysis as they both use computational experimentation for the treatment of uncertainty. However, contrary to sensitivity analysis, exploratory modelling does not have any pre-assumptions regarding the probability distribution of the uncertainty space. In other words, exploratory modelling can deal with Knightian form of uncertainty where no information (for example, ranking, probability distributions, estimates) exist about the uncertain parameters.

Experimental design

Exploratory modelling is based on the generation and analysis of computational experiments. The way that computational experiments are established and performed—that is, their experimental design— influences the nature and number of generated results, the insights gained from them, as well as the computational cost and time to complete the experiments (Kwakkel and Pruyt 2015). Experimental design includes a decision regarding the list of critical uncertainties, the space of the uncertainty, a sampling strategy for choosing from this space, an appropriate number of samples, and the response variables (outcomes of interest) in the experiments. Among them, delineating the uncertainty space is a delicate aspect of experimental design. An extreme uncertainty space can lead to an extreme computational burden and too plural (many) conclusions, and a very narrow uncertainty space can lead to the risk of missing potential future possibilities from the analysis. Several approaches have been introduced to delineate the uncertainty space, including: setting independent uniform distributions with lower and upper bounds (Pianosi *et al*. 2016), limiting the feasible uncertainty space using *a priori* knowledge for filtering subsets associated with a certain outcome (Kasprzyk *et al*. 2013), and assigning likelihood weights to different values from the uncertainty space based on a comparison between model-generated and observed values (Beven and Binley 1992).

METHODS

We use four methods for answering the two questions of this article. This section introduces these methods briefly. The way that we use each method is explained later.

ANOVA

ANalysis Of VAriance (ANOVA) is a statistical technique to test the statistical significance of difference between means of multiple groups (Montgomery 2001). ANOVA is based on a null hypothesis of no significant difference among groups and an alternative hypothesis of at least one significant difference among groups. Assuming initially that the null hypothesis is true, the observed difference of group means is called statistically significant if it is concluded unlikely to happen by chance. ANOVA uses the F-statistic (that is, a ratio of two variances) to test the statistical difference. ANOVA generate the F-statistic and compared its associated probability of occurrence (p-value) with a threshold (significance level). A p-value less than a threshold justifies the rejection of the null hypothesis and the support of the alternative hypothesis. See Iverson and Norpoth (1987) for further explanation of this technique.

Multi-dimensional clustering

Multi-dimensional clustering (Gerst *et al*. 2013) groups many potential model behaviours, generated based on different samples from the input uncertainty space, into clusters of similar behaviours. The appropriate number of clusters is decided based on the value of Bayesian information criterion (BIC) and Aikake's information criterion (AIC) (McLachlan and Peel 2004). Multi-dimensional clustering uses this appropriate number of clusters and generates a mixture of Gaussian distributions to estimate the distribution of system behaviour in the chosen number of clusters.

Scenario discovery

Scenario discovery is a statistical, data-mining process used to identify subsets of the input uncertainty space that result in similar classes of behaviour in the output solution space (Bryant and Lempert 2010, and Groves and Lempert 2007). Scenario discovery starts by generating computational experiments using a model based on input samples from the input uncertainty space. It distinguishes similar classes of behaviour among experiments and selects alternative subsets from the input uncertainty space (in the format of hyper-dimensional boxes) to describe the classes of behaviour. Scenario discovery uses two measures of quality (coverage and density) and a p-value for comparing the generated subsets and for choosing the best subset of the uncertainty space responsible for the creation of the behaviour of interest in the model. Coverage describes how universally a subset can cover all experiments from a same class of behaviour, and density describes to which extent samples from a subset can only result in experiments with a certain class of behaviour and no other behaviours. Scenario discovery has been implemented using a number of algorithms (Lempert *et al*. 2008) such as classification and regression tree (CART) (Breiman et al. 1984) and patient rule induction method (PRIM) (Friedman and Fisher 1999). See (Moallemi *et al*. 2017, Lempert *et al*. 2013, Lempert and Groves 2010, and Moallemi and Malekpour 2018) for implementations of scenario discovery.

Multi-objective robust optimisation

Multi-objective robust optimisation is a group of methods which generate alternative solutions for maximising or minimising multiple objective functions under constraints. The solutions much remain valid under any future conditions (Deb 2001, Marler and Arora 2004, and Ben-Tal and Nemirovski 2000). The result of multi-objective robust optimisation is not a single optimal solution. The multiplicity of (conflicting) objectives necessitates the generation of *Pareto optimal* solutions, which can compromise among multiple objectives. The presence of future uncertainties also necessitates *robust* solutions where the performance of solutions remains insensitive to drastic changes in the input parameters in the face of future uncertainties. Multi-objective evolutionary algorithms such as NSGAII (Deb and Pratap 2002) and Borg (Hadka and Reed 2013), also see Maier *et al.* (2014) and the non-dominated sorting search algorithm such as Woodruff and Herman (2013) are among popular methods for generating and post-processing (respectively) of Pareto and robust solutions in multiobjective optimisation problems. In both cases, a simulation model is used to assess the impact of the solution space on objective functions and to identify the Pareto robust solutions. See (Moallemi *et al.* Submitted-b, Kasprzyk *et al.* 2013, and Hamarat et al. 2014) for the implementations of multiobjective robust optimisation.

Table 1. List of uncertain parameters with their ranges of variation

Uncertain parameter	Range
The risk that an aircraft is lost during operation	0.00026 – 0.00234 (–)
Lifetime of aircraft	37440 – 336690 (hour)
Total required flying hours	12 – 200 (hour/week)
Expected time spent by an aircraft in CAP	8 – 45 (week)
Time between CAP events	16 – 40 (week)
Expected time spent by an aircraft in DM (Time in DM)	5 – 25 (week)
Time (flying hours) between DM events	200 – 1800 (hour)
Expected time spent by an aircraft in OM (Time in OM)	3 – 15 (week)
Time between OM events	50 – 450 (hour)
Available capacity for CAP	1 – 7 (–)
Number of purchased aircrafts	1 – 7 (–)
OM available capacity	1 – 7 (–)
DM available capacity	1 – 7 (–)

Table 2. The decision space

Decision variable	Range
Available capacity for CAP	1 – 7 (–)
Number of purchased aircrafts	1 – 7 (–)
OM available capacity	1 – 7 (–)
DM available capacity	1 – 7 (–)

THE DELINEATION OF THE INPUT UNCERTAINTY SPACE IN EXPLORATORY MODELLING

This section addresses the two questions raised in Introduction using an illustrative example of asset acquisition and management of aircraft fleets.

THE SIGNIFICANCE OF THE INPUT UNCERTAINTY SPACE

This section shows the significance of the way we delineate the input uncertainty space. While it is clear that changing the delineation of uncertainties impacts the results, this section aims to show the statistical significance of this impact and also to demonstrate the variation in results visually. We explain the steps taken, the generated results and discussion as follows.

Process

We assess the significance of the input uncertainty space in three steps based on the methods explained in Methods:

- Step 1: Different ensembles of computational experiments are generated based on sampling from full and truncated input uncertainty spaces.
- Step 2: A joint kernel density estimate (KDE) diagram is used to show the variation the output solution space — in terms of the state of selected model outputs — in response to the subsetting of the input uncertainty space. ANOVA is also used to test the statistical significance of variation among the multiple ensembles of generated experiments with full and truncated input uncertainty space.
- Step 3: The distributions of the output solution space — in terms of states of selected model outputs in the last time step of the simulation duration — are represented in series of scatter plots, each plot based on sampling from one specific area of the input uncertainty space. The states of the model outputs in the solution space are also clustered using the multi-dimensional clustering technique. We compare the clusters and scatter plots to show how the choice of delineation in the input uncertainty spaces impacts the distribution of the output solution space.

Results and discussion

To analyse the impact of sampling from different areas of the input uncertainty space, we generated three ensembles of experiments, each including 200 model runs. Ensembles are generated based on sampling from the full, the first quartile, and the fourth quartile of the uncertainty space and the decision space—where different combinations of decisions regarding the size of acquisition and maintenance can be chosen from. See Table 1 for the full range of uncertainties (also see Appendix A). See Table 2 for the range of decision variables. Experiments in each ensemble represent the model response in terms of (average) in-service aircraft and total (acquisition and maintenance) costs, as two outcomes of interest.

The distribution of the output solution space—in terms of the state of in-service aircraft and total costs—in each ensemble was drawn in a joint KDE diagram (see Figure 1). It is observed that the choice of the input uncertainty space in each ensemble creates a different distribution of the solution space. This difference is more visible in the distribution of in-service aircraft (Figure 1 (b)). This observation can prove our initial assumption regarding the significance of the delineation of the input uncertainty space. We also performed ANOVA to statistically verify the visual observation.

The result of ANOVA in Table 3 shows that in both model outcomes, the null hypothesis (similarity of the means of distributions) is rejected and the alternative hypothesis is supported.

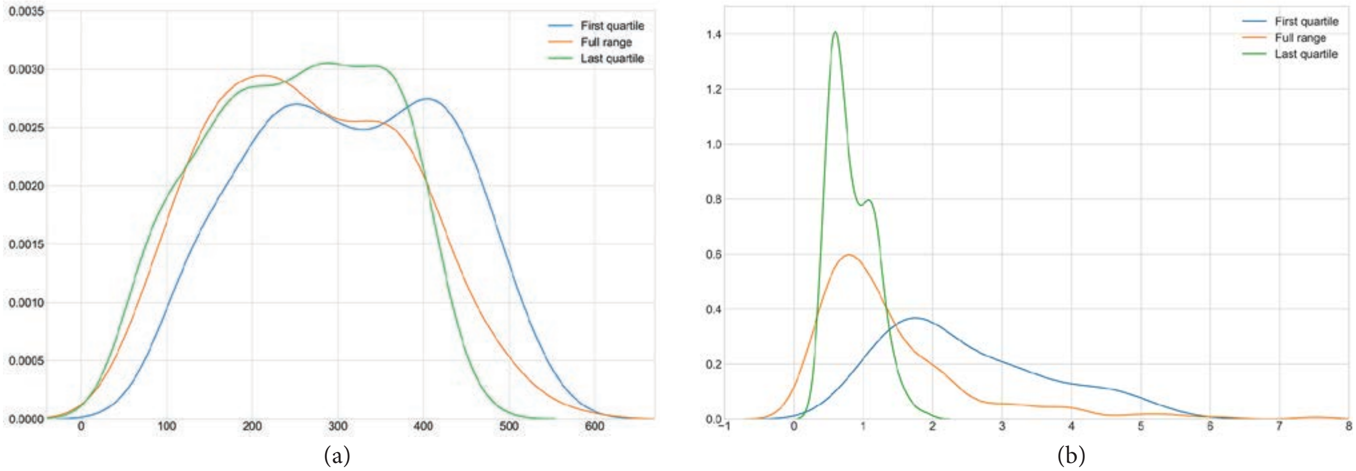


Figure 1. KDEs for (a) total costs and (b) in-service aircraft in three ensembles of experiments for the full range, first quartile, and fourth quartile of uncertain parameters

Table 3. The results of ANOVA (5% significance level) for the thhree ensembles of experiments for (a) total costs and (b) in-service aircraft

SUMMARY				
Ensemble	**Count**	**Sum**	**Average**	**Variance**
Full range	200	53070	265.35	12574.16
First quartile	200	62661	313.305	12917.03
Last quartile	200	50184	250.92	10148.98

ANOVA		
F	**P-value**	**F critical**
17.95669	2.67E-08	3.010815

(a)

SUMMARY				
Ensemble	**Count**	**Sum**	**Average**	**Variance**
Full range	200	278.6238	1.393119	1.323377
First quartile	200	496.2475	2.481238	1.387502
Last quartile	200	166.5347	0.832673	0.096493

ANOVA		
F	**P-value**	**F critical**
150.1713	1.48E53	3.010815

(b)

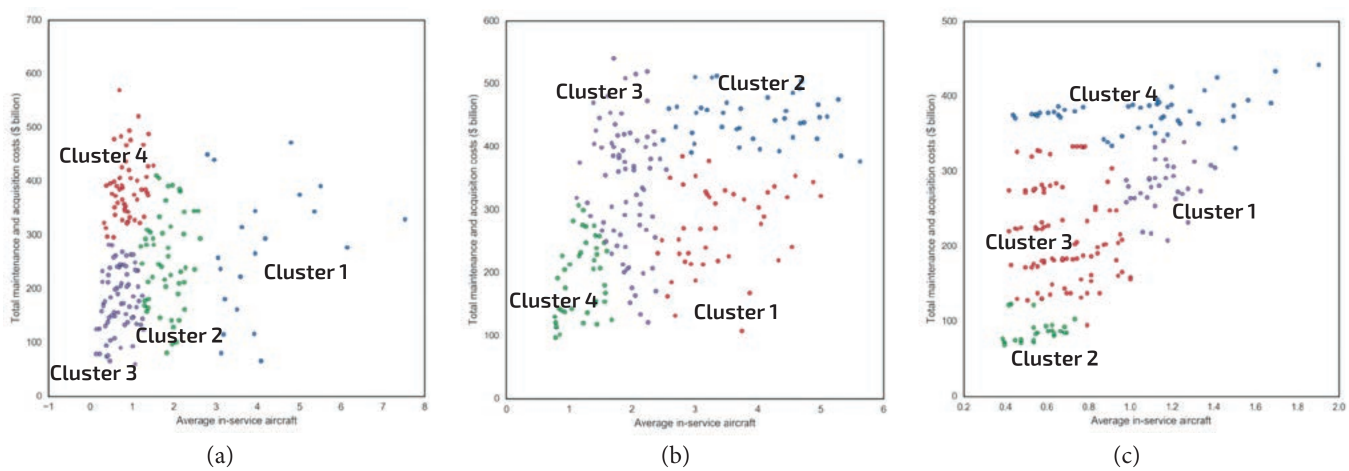
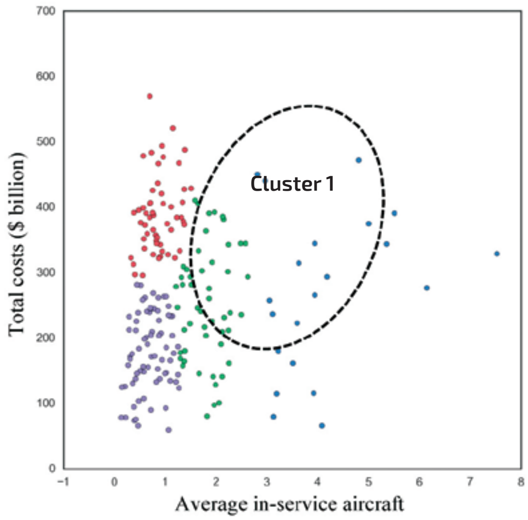


Figure 2. Clusters of experiments with similar behaviour regarding in-service aircraft and total costs based on (a) the full range, (b) first quartile, and (c) fourth quartile of uncertain parameters. Note that clusters are named randomly and the features of clusters from one plot to another do not remain similar.

Solution space



Input uncertainty space

The truncated uncertainty ranges. For the rest of uncertainties in Table 1, the full ranges are considered.

Measure of merits	Uncertainty	Range	P-value
Coverage: 0.28	Required flying hours	13–61 (hours/week)	7.0e–05
Density: 1.00	DM available capacity	4–6 (–)	4.6e–02
Last quartile	200	166.5347	0.832673

Figure 3. The relationship between Cluster 1 in the solution space and the input uncertainty space

We clustered and plotted generated experiments in each ensemble with respect to the state of in-service aircraft and total costs in a scatter plot (see Figure 2). The results show that truncated uncertainty spaces result in a high resolution of the solution space, but they also disregard some possible variations in the output solution space. For example, Figure 2 (a), based on the full range of uncertainties, shows the possible variations in the solution space (i.e., from about 0 to 8 in-service aircraft and from about B$50 to B$600 total costs) in a high-level picture. However, Figure 2 (b) and Figure 2 (c), which are based on truncated ranges of uncertainties, dismiss this wide possible variation of in-service aircraft and instead present a higher resolution and more accurate clustering of the solution space in regions with high (that is, 1 to 6) in-service aircraft and low (that is, 0.4 to 2) in-service aircraft respectively. A high resolution and accurate clustering of the solution space and the inclusion of a wide diversity of variations can be achieved together if we consider the full uncertainty space and also increase the number of samples (experiments) at the same time. However, this can come at the costs of adding to the computation burden of the process. This leads us to conclude that there is a need for a smart approach which can truncate the input uncertainty space effectively; in a way that improves the resolution and accuracy of the output solution space while minimising the exclusion of relevant solution possibilities from this space. We introduce a heuristic to address this need in the next section.

To answer the first question posed in Introduction, the output solution space is sensitive to changes in the input uncertainty space as full and truncated areas of the input uncertainty create different (statistically significant) means and dispersion for the distribution of the output solution space.

A HEURISTIC TO INFORM THE DELINEATION OF THE INPUT UNCERTAINTY SPACE

The previous section demonstrated the implications of sampling from the different areas of the input uncertainty space for the output solution space. This section presents a heuristic to inform the delineation of the uncertainty space. The idea behind this heuristic is that although exploratory modelling can generate a wide solution space, not all solutions in this space are relevant for the context of study. These solutions are technically possible to be generated based on random samples from the input uncertainty space. However, the solutions are not possible in reality or, even if they happen, they are not intended by decision-makers or are not relevant to the context. Stakeholders, with their practical knowledge of the context, can help to identify these irrelevant parts of the solution space. We then sample only from those areas of the input uncertainty space which can generate the relevant parts of the solution space. We explain the steps taken in this heuristic, the generated results, and discussion as follows.

Process

To answer the second question posed in Introduction, we suggest a heuristically informed data-mining process which modifies the input uncertainty space based on screening the output solution space in the following steps:

- Step 1: The process starts by projecting the clusters of similar parts of the output solution space (in terms of selected model outcomes) in potential futures with a data-mining technique, called multi-dimensional clustering.
- Step 2: Stakeholders are asked about the desired/relevant parts of this solution space. This stakeholder opinion is used as a heuristic to limit the solution space.
- Step 3: A data-mining method, called scenario discovery, is used to identify which regions of the input uncertainty space are more likely to be responsible for the generation of the desired/relevant parts of the solution space.

Results and discussion

To demonstrate the implementation of this process, we plotted the potential solution space (in terms of in-service aircraft and total costs) based on the full ranges of uncertainties in a scatter plot (see Figure 3). We assumed that the behaviour of experiments in Cluster 1, which are featured with higher in-service aircraft, is of more interest to decision-makers and relevance to the context compared to the other clusters. We, therefore, discard other clusters and only focus on sampling from the area of the input uncertainty space responsible for the generation of experiments in Cluster 1. To delineate this specific area of the uncertainty space, we applied scenario discovery.

The results of the scenario discovery show that the input uncertainty space would be more likely to result in Cluster 1 if the required flying hours and DM capacity available (from Table 1) are truncated to the specified ranges in Figure 3 while the

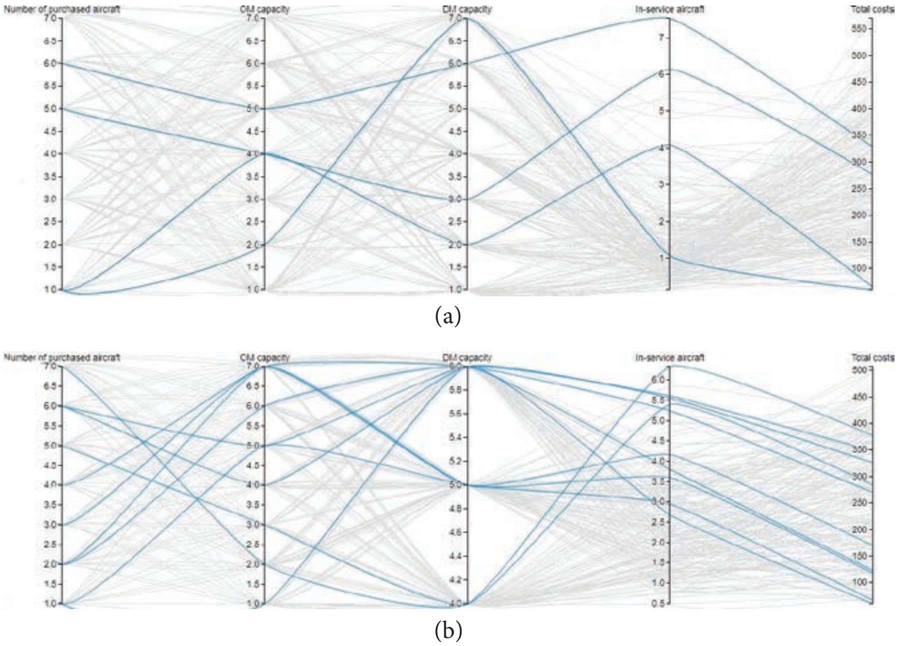


Figure 4. Pareto optimal solutions based on sampling from (a) the full input uncertainty space and (b) the truncated uncertainty space

ranges of the rest of uncertainties in Table 1 remain the same. This truncated input uncertainty space improves the efficiency of the exploration process by avoiding having to sample from the areas of the input space insensitive to the desired solutions.

We show the implication of our heuristic approach for the exploratory modelling process with a multi-objective decision-making problem aiming to find the number of purchased aircraft, OM capacity and DM capacity, which could maximise in-service aircraft and minimise total costs. We generated Pareto optimal solutions with multi-objective robust optimisation under two conditions: the full input uncertainty space and the uncertainty space truncated according to the results of scenario discovery (in Figure 3). This resulted in two different sets of solutions in Figure 4. One reason for their difference is that solutions under each condition are generated based on two series of random sampling and simulation runs. However, the more insightful reason of the difference lies in the different areas of the input uncertainty on which the generation solutions are based. The comparison of the Paretooptimal solutions in two conditions (see Figure 4) shows that running multi-objective robust optimisation under an informed-delineated input uncertainty space can result in more solutions with desired performance for decision-makers. The solutions in Figure 4 (b) can deliver a better trade-off among the multiple decision objectives by delivering a higher in-service aircraft and not necessarily resulting in a higher total cost compared to solutions in Figure 4 (a).

CONCLUSIONS

The delineation of an appropriate uncertainty space in the exploratory modelling process is significant as it needs to capture a picture of model behaviours that is as wide as possible bit does not incur a high computational cost. This led us to think about the ways that we can effectively adjust the delineation of the input uncertainty space based on a feedback control from output solutions. We argued that stakeholders can screen the solution space and identify the desired behaviour of the model. A feedback control, then, can identify which area of the input uncertainty space is responsible for the generation of the desired model behaviour and can inform the adjustment of input uncertainty space accordingly.

The feedback control that we suggested in this article is based on data mining and statistical analyses of both input uncertainties and output solutions. However, this is not the only way to design such feedback control. An alternative approach would be to develop a simple form of control model which can relate outputs to inputs. This control model is much simpler than the original simulation model which generated the outputs from inputs in the first place. Therefore, the control model can be run very quickly to inform the delineation of the input uncertainty space as the desired model behaviour is identified in the solution space. Although this control model would not be accurate in explaining the precise input-output relationship because of its simple structure, the accuracy of the control model can be improved by training the model with many ensembles of inputs and outputs. We suggest the development of this control model in the exploratory modelling process and its comparison with our suggested control data mining and statistical analyses as a future research direction. ■

REFERENCES

- Bankes, S. 1993. “Exploratory modeling for policy analysis.” *Operations Research* 41: 435-449.
- Bankes, S. C., R. J. Lempert, and S. W. Popper. 2001. “Computer-assisted reasoning.” *Computing in Science & Engineering* 3: 71-77.
- Ben-Haim, Y. 2006. *Info-Gap Decision Theory: Decisions Under Severe Uncertainty*, London, GB: Academic Press.
- Ben-Tal, A. and A. Nemirovski. 2000. “Robust solutions of linear programming problems contaminated with uncertain data.” *Math. Program*. 88: 411.
- Bergmann, S., S. Stelzer, and S. Straßburger. 2011. *Initialization of simulation models using CMSD. In Proceedings of the 2011 Winter Simulation Conference (WSC)*, pp. 2223-2234. IEEE.
- Beven, K., and A. Binley. 1992. “The future of distributed models: model calibration and uncertainty prediction.” *Hydrological processes* 6: 279-298.
- Breiman, L., J. Friedman, C. J. Stone, and R. A. Olshen. 1984. *Classification and regression trees*. CRC Press.
- Bryant, B. P. and R. J. Lempert. 2010. “Thinking inside the box: A participatory, computer-assisted approach to scenario discovery.” *Technological Forecasting and Social Change* 77: 34-49.
- Deb, K. 2001. *Multi-objective optimization using evolutionary algorithms*. John Wiley & Sons.
- Deb, K., and A. Pratap. 2002. “A fast and elitist multiobjective genetic algorithm: NSGA-II.” *IEEE Trans. Evol. Comput*. 6: 182.
- Friedman, J. H., and N. I. Fisher. 1999. “Bump hunting in high-dimensional data.” *Statistics and Computing* 9: 123-143.
- Gerst, M. D., P. Wang, and M. E. Borsuk. 2013. “Discovering plausible energy and economic futures under global change using multidimensional scenario discovery.” *Environmental Modelling & Software* 44: 76-86.

- Groves, D. G., and R. J. Lempert. 2007. "A new analytic method for finding policy-relevant scenarios." *Global Environmental Change* 17: 73-85.
- Haasnoot, M., J. H. Kwakkel, W. E. Walker, and J. ter Maat. 2013. "Dynamic adaptive policy pathways: A method for crafting robust decisions for a deeply uncertain world." *Global Environmental Change* 23: 485-498.
- Hadka, D. and P. Reed. 2013. "Borg: An auto-adaptive many-objective evolutionary computing framework." *Evol. Comput.* 21: 231.
- Hamarat, C., J. H. Kwakkel, E. Pruyt, and E. T. Loonen. 2014. "An exploratory approach for adaptive policymaking by using multi-objective robust optimization." *Simulation Modelling Practice and Theory* 46: 25-39.
- Iverson, G. R. and H. Norpoth. 1987. *Analysis of variance*, Thousand Oaks, US-CA: Sage Publications.
- Kasprzyk, J. R., S. Nataraj, P. M. Reed, and R. J. Lempert. 2013. "Many objective robust decision making for complex environmental systems undergoing change." *Environmental Modelling & Software* 42: 55-71.
- Kwakkel, J. H. 2017. "The Exploratory Modeling Workbench: An open source toolkit for exploratory modeling, scenario discovery, and (multi-objective) robust decision making." *Environmental Modelling & Software* 96: 239-250.
- Kwakkel, J. H. and E. Pruyt. 2015. "Using System Dynamics for Grand Challenges: The ESDMA Approach." *Systems Research and Behavioral Science* 32: 358-375.
- Kwakkel, J. H., W. E. Walker, and V. A. W. J. Marchau. 2010. "Adaptive airport strategic planning." *European Journal of Transport and Infrastructure Research* 10: 249–273.
- Lempert, R. 2013. "Scenarios that illuminate vulnerabilities and robust responses." *Climatic Change* 117: 627-646.
- Lempert, R. J., B. P. Bryant, and S. C. Bankes. 2008. "Comparing algorithms for scenario discovery." RAND, Santa Monica, US-CA.
- Lempert, R. J. and D. G. Groves. 2010. "Identifying and evaluating robust adaptive policy responses to climate change for water management agencies in the American west." *Technol. Forecast. Soc.* 77: 960.
- Lempert, R. J., N. Kalra, S. Peyraud, Z. Mao, S. B. Tan, D. Cira, and A. Lotsch. 2013. Ensuring Robust Flood Risk Management in Ho Chi Minh City. *Policy Research Working Paper 6465*. World Bank.
- Lempert, R. J., S. W. Popper, and S. C. Bankes. 2003. *Shaping the next one hundred years: new methods for quantitative, long-term policy analysis*. RAND Corporation.
- Maier, H. R., Z. Kapelan, J. Kasprzyk, J. Kollat, L. Matott, M. Cunha, G. C. Dandy, M. S. Gibbs, E. Keedwell, and A. Marchi. 2014. "Evolutionary algorithms and other metaheuristics in water resources: Current status, research challenges and future directions." *Environmental Modelling & Software* 62: 271-299.
- Marler, R. T., and J. S. Arora. 2004. "Survey of multi-objective optimization methods for engineering." *Structural and multidisciplinary optimization* 26: 369-395.
- McLachlan, G., and D. Peel. 2004. *Finite mixture models*. John Wiley & Sons.
- Moallemi, E. A., F. de Haan, J. Kwakkel, and L. Aye. 2017. "Narrative-informed exploratory analysis of energy transition pathways: A case study of India's electricity sector." *Energy Policy* 110: 271–287.
- Moallemi, E. A., S. Elsawah, and M. J. Ryan, M. J. Submitted-a. "Robust Decision Making and Epoch-Era Analysis: A comparison of two model-based approaches for decision making under uncertainty." *Technological Forecasting & Social Change*.
- Moallemi, E. A., S. Elsawah, and M. J. Ryan. Submitted-b. "Model-based multi-objective decision making under deep uncertainty from a multi-method design lens." *Simulation Modelling Practice & Theory*.
- Moallemi, E. A., and S. Malekpour. 2018. "A participatory exploratory modelling approach for long-term planning in energy transitions." *Energy Research & Social Science* 35: 205-216.
- Montgomery, D. C. 2001. *Design and Analysis of Experiments*. New York: Wiley.
- Pianosi, F., K. Beven, J. Freer, J. W. Hall, J. Rougier, D. B. Stephenson, and T. Wagener. 2016. "Sensitivity analysis of environmental models: A systematic review with practical workflow." *Environmental Modelling & Software* 79: 214-232.
- Schulze, T., S. Strassburger, and U. Klein. 1999. "On-line data processing in simulation models: New approaches and possibilities through HLA." In *Proceedings of the 31st conference on Winter simulation: Simulation—a bridge to the future-Volume 2*, pp. 1602-1609.
- Walker, W. E., M. Haasnoot, and J. H. Kwakkel. 2013. "Adapt or perish: a review of planning approaches for adaptation under deep uncertainty." *Sustainability* 5: 955-979.
- Woodruff, M., and J. Herman. 2013. *Nondominated sorting for multi-objective problems* [Online]. Available: https://github.com/matthewjwoodruff/pareto.py [Accessed 2017].

ABOUT THE AUTHORS

[Editor: Author biographies were current when the paper was initially published in 2018.]

Enayat A. Moallemi is research associate at the Capability Systems Centre, the University of New South Wales, Canberra. His research lies at the intersection of robust decision making, exploratory modelling, and sustainability transitions. He specifically explores the long-term behaviour of complex systems under deep uncertainty for informing robust decision making.

Sondoss Elsawah is a senior lecturer in the School of Engineering and IT (SEIT) and leads the modelling and simulation effort in the Capability Systems Centre. Dr Elsawah's research program focuses on understanding the behaviour of large complex problems and systemic risks that arise from the interactions between social, ecological, and technological systems. This focus includes how to use that understanding to create computational tools that enable decision makers design and evaluate options to deal with such complexity.

Michael J Ryan is associate professor and the director of the Capability Systems Centre. In addition, he is the program coordinator for the Master of systems engineering program in the School of Engineering and Information Technology (SEIT). He is the author or co-author of twelve books, three book chapters, and over 200 technical papers and reports.

APPENDIX A on following page

Appendix A

Table A.1. The first quartile of uncertainty ranges

Uncertain parameter	Range
The risk that an aircraft is lost during operation	0.00026 – 0.00078 (–)
Lifetime of aircraft	37440 – 112252.5 (hour)
Total required flying hours	12 – 59 (hour/week)
Expected time spent by an aircraft in CAP	8 – 17.25 (week)
Time between CAP events	16 – 22 (week)
Expected time spent by an aircraft in DM (Time in DM)	5 – 10 (week)
Time (flying hours) between DM events	200 – 600 (hour)
Expected time spent by an aircraft in OM (Time in OM)	3 – 6 (week)
Time between OM events	50 – 150 (hour)

Table A.2. The fourth quartile of uncertainty ranges

Uncertain parameter	Range
The risk that an aircraft is lost during operation	0.00182 – 0.00234 (–)
Lifetime of aircraft	261877.5 – 336690 (hr)
Total required flying hours	153 – 200 (hour/week)
Expected time spent by an aircraft in CAP	35.75 – 45 (week)
Time between CAP events	34 – 40 (week)
Expected time spent by an aircraft in DM (Time in DM)	20 – 25 (week)
Time (flying hours) between DM events	1400 – 1800 (hour)
Expected time spent by an aircraft in OM (Time in OM)	12 – 15 (week)
Time between OM events	350 – 450 (hour)

Flanigan and Dixon continued from page 22

Jeffery Dixon is a senior combat effectiveness analyst with 17 years of experience in the design and analysis of future combat systems. Jeff's main area of expertise is modeling, simulation, and analysis using design of experiments principles. Prior to joining JHU/APL in 2001 Jeff served in the US Surface Navy as a missile officer. He also teaches the course Metrics for Modeling and Simulation in the JHU systems engineering curriculum. He is currently employed by The Johns Hopkins University Applied Physics Laboratory as the Tactical Aircraft Platforms Program Manager.

Hamida, Jankovic, Huet and **Bocquet** continued from page 30

Sonia Ben Hamida has been working as a system engineer since 2011 at Airbus Safran Launchers (ASL), the European leader in space transportation. First, she took part in the Single European Sky Air traffic management Research (SESAR) program for 2 years. Then, she started a PhD in partnership with the SystemX research institute and the *École CentraleSupélec* – one of the top French research institutes in complex system design. Her research focuses on how to design value proposition for new businesses under uncertainty. Sonia holds an engineer's degree in aerospace from the French Civil Aviation *grande école*. She is a member of the French INCOSE chapter and the Design Society.

Marija Jankovic is an associate professor at CentraleSupélec, Université de Paris Saclay. Her main domain of interest concerns developing a decision support framework for early design stages. She is interested in developing support methods and tools that will permit design engineers to make more robust decisions. Her research work is also challenged by multidisciplinary design environments that are developing in view of the new world's competition. Dr. Jankovic has working experience in designing complex systems. Most of her research projects are performed in collaboration with industry or government and with direct implementation and verification of research results. She collaborates with some of the major French and international companies such as Snecma, Thales, Airbus, PSA Peugeot Citroen, and Schlumberger.

Alain Huet works with Airbus Safran Launchers. He is the head of the Complex Systems Architecture department, in charge of systems engineering. He worked 3 years on the European program SESAR (Single European Sky Air traffic management Research) on requirements management and verification and validation. Before, he was the head of the ground systems department, dedicated to the design of qualification facilities and simulators for the ATV (automated transfer vehicle). He has also been involved in systems of systems programs for NATO. Alain Huet graduated from the Rennes University of Technology.

Jean-Claude Bocquet: engineering design methodology for complex systems, industrial engineering academic background full professorship qualifications in mechanical engineering and automatic, since 1990, at Ecole Centrale Paris "Habilitation à Diriger des Recherches" (university professor degree in France), Nancy University, France, 1990, agrege professor, mechanical engineering design, Ecole Normale Superieure de Cachan, 1977–1987. PhD in sciences, Université Pierre et Marie Curie Paris VI, 1981, masters (DEA) in mechanical engineering, University Pierre et Marie Curie Paris VI, 1977 agregation in mechanical engineering, ranking 2nd, Ecole Normale Supérieur de Cachan, 1976, professional experience head of industrial engineering research laboratory of Ecole Centrale Paris (55 PhD students) Laboratoire Génie Industriel http://www.lgi.ecp.fr. Head of Enterprise Sciences Department of Ecole Centrale Paris, member and vice president of the Board of Ecole Centrale Paris 2004-2013. Expert member, Agence Nationale de la Recherche (ANR), Expert member Agence d'Evaluation de la Recherche et de l'Enseignement Supérieur (AERES), Expert member of Industry Ministry.

Assessing the Impacts of Uncertainty Propagation to System Requirements by Evaluating Requirement Connectivity

Alejandro Salado and **Roshanak Nilchiani**

ABSTRACT
Although theoretically independent, requirements within a decomposition level of a system architecture are not isolated elements. For an existing design, a change of a requirement may endanger or facilitate fulfillment of other requirements within the same level of the decomposition. The present research suggests a requirement connectivity metric to evaluate the potential consequences that changing a requirement may have on a system with respect to fulfillment of other requirements. A particular aspect of the present research is the assumption that connectivity accounts only for requirements within the same decomposition level of an architecture, not for those flowing up or down the decomposition. The metric is used to evaluate different cases in which requirements are changed due to triggering of uncertain events during a project life cycle.

INTRODUCTION

It is widely recognized that various types of change during a development process often negatively affect cost, schedule, and quality. Consequently, the study of how change propagates along the elements composing a system that is under development is becoming an important topic of research in different engineering disciplines and industries. For example, (Hassan and Holt 2004) address the importance of predicting change in software systems and provide some heuristics for its management. Heuristics are also proposed in the field of product development (Keller, Eckert, and Clarkson 2006), where the majority of research has focused attention to the impacts on physical components derived from changes on other physical components. Clarkson, Simons, and Eckert (2001) and Oduncuoglu and Thomson (2011) include risk as a measure on the prediction of change propagation of product development. In research spanning several years and over 41,500 change requests in a number of industry cases (Giffin et al. 2009) provide some metrics to quantitatively evaluate the effect of change propagation and find out that the majority of change requests occur during system integration and test.

Changes can however occur or be a result of changes in elements or artifacts surrounding the system under development, apart from the elements of the system under study. Koh and Clarkson (2009) explore this idea and incorporate other elements in the evaluation of change propagation during product development, namely design features and requirements (which in systems engineering jargon can be interpreted as stakeholder requirements and system requirements). However, the authors do not address interdependencies between stakeholder requirements, but only between these and system requirements. Bonjour et al. (2010) provide a new hindsight to the state of the art and address the effects project organizational aspects have on change propagation.

In the development of complex systems, considerable number of changes are the result of requirement changes. Peteerson et al. (2007) discuss the effects that requirement changes have in projects and propose some guidelines to deal with such uncertainty in multidisciplinary innovation products. Bischof and Blessing (2007) explore the same path to discover how to design products for flexibility so that they can seamlessly absorb such changes and argue that the more frequent requirements change, the more decisions in shorter time need to be taken. Consequently, the uncertainty of taking good decisions increases, which could thus result in not meeting customer expectations and having to redesign the product. Such a concern is also addressed by Sudin and Ahmed-Kristensen (2011), who investigate how requirement change occurs in the context of a design process. Eben and Lidemann (2010) use the findings of change propagation in product development to evaluate change propagation in requirements and recognize interdependencies exist among requirements within a decomposition level.

Hull, Jackson, and Dick (2005) state that "a simple textual statement is not sufficient fully to define a requirement; there is other classification and status information that each requirement carries." They propose that such information should be attached to requirements as attributes and provide a comprehensive list of 34 attributes organized in 9 categories, which are in part, taken from the work carried out by the requirements working group in the UK chapter of INCOSE. One of such attributes is priority, which by itself may represent also multiple objectives. Firesmith (2004) justifies the importance of prioritizing requirements and shows the broadness of meaning it can convey by collecting types of priorities that are used or have been used within industry and academia: stakeholder preferences, value to business, cost, harm avoidance, difficulty to implement, and connectivity, among others. Carlshamere et al. (2001) address the importance of managing requirement dependencies as an enabler to establish effective prioritization techniques. Kulshreshtha, Boardman, and Verma (2012) identify the same problematic when analyzing prioritization based on pair-wise comparisons as proposed by Karlsson (1996), Mead (2006), and Berander and Andrews (2005).

Connectivity is defined in the present research as the number of requirements within a decomposition level of a system and for a single element in the decomposition level to which a particular requirement has a relation (Figure 1). Consequently, it indicates how many requirements within a decomposition level may see their fulfillment affected due to a change on a particular requirement. In essence, connectivity in the present research measures the potential harm or effect that a requirement change may have on the fulfillment of other requirements by the system.

The present research contributes to the field of change propagation by incorporating uncertainty as the driver for connectivity evaluation (thus pro-actively contributes to design against change versus the reactive approach to changes), proposing a connectivity metric based on the influence impact of a requirement, and incorporating structural definitions to links and nodes concurrently.

The present paper is organized as follows. First a literature survey is presented on the topic of requirement connectivity. Then a case study and the connectivity metrics that are used in the research are presented. The paper then continues with the evaluation of 10 cases for which the impact of triggered uncertain events on requirements showcases the use of requirement connectivity. Finally, a short summary of the conclusions and a proposal for future work are given.

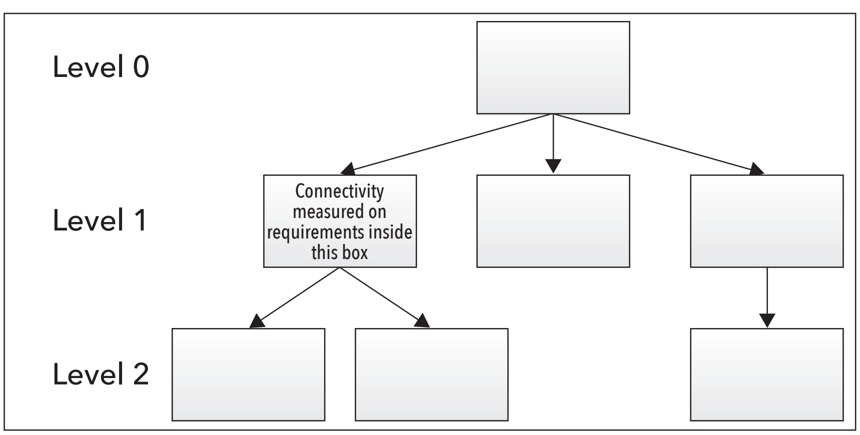


Figure 1. Definition contest for connectivity in the present research

LITERATURE REVIEW

The issue of interdependencies between requirements has been a topic of interest in recent years in industry as well as academia. Kulshreshtha, Boardman, and Verma (2012) provide a comprehensive literature review on the different propositions for defining dependencies among requirements, which is summarized hereafter.

Carlshamere et al. (2001) in research using industrial data, define the following six categories of requirement dependence, although they recognize that not all relationships they found could be categorized with it:

- *AND*: indicates a bidirectional dependency in which one or more requirements need the other ones to be fulfilled.
- *REQUIRES*: indicates a unidirectional dependency in which one requirement needs another one to be fulfilled.
- *TEMPORAL*: indicates a requirement needs another one to be fulfilled before it can be fulfilled itself.
- *CVALUE*: fulfillment of a requirement affects fulfillment of a cost requirement.
- *ICOST*: fulfillment of a requirement affects fulfillment of a time requirement.
- *OR*: a requirement does not need to be fulfilled if another one is fulfilled.

Pohl (1996) uses traceability to propose 18 types of dependencies classified in 5 categories (the names are self-explanatory):

- *Condition*: Constraints, Pre-conditions.
- *Content*: Similar, Compares, Contradicts, Conflict.
- *Documents*: Example for, Test_case for, Purpose, Background, Comments.
- *Evolutionary*: Elaborates, Formalizes, Based_on, Satisfies, Replaces.
- *Abstraction*: Refines, Generalizes.

Zhan, Mei, and Zhao (2005) take a different approach and focus on dependencies of features, grounded in the hypothesis that features are the outcomes of subsets of requirements and in the fact that previous classifications did not address requirements dependencies in the solution space. They propose the classification shown in Table 1.

Table 1. Requirement dependency taxonomy (Zhan, Mei, and Zhao 2005)

Type	Description
Static	
Decomposition	Feature flow relations (originating and derivation).
Generalisation	Higher level feature identified from a lower level one.
Required or excluded	Unidirectional dependence or exclusion.
Dynamic	
Serial	Pre-condition.
Collateral	Simultaneous bi-directional dependency.
Synergetic	Serial bi-directional dependency.
Change	Change in a feature affects another one

Robinson, Pawloaski, and Volkov (1999)

propose 9 types of dependencies which aim at supporting interaction management in conflicting situations:

- *Positive correlation*: increasing the fulfillment of a requirement increases the fulfillment of another requirement.
- *Negative correlation*: increasing the fulfillment of a requirement decreases the fulfillment of another requirement.
- *Unspecified correlation*: a change in the fulfillment of a requirement has an unidentified effect on another requirement.
- *No correlation*: change in the fulfillment of a requirement does not impact the fulfillment of another requirement.
- *Structure*: requirements are similar.
- *Resource*: requirements depend on the same resource.
- *Task*: a requirement describes a dependent task of another requirement.
- *Causality*: a requirement is a consequence of another requirement.

Table 2. Example requirements

ID	Requirement
Req1	The system shall image the Earth surface in 4 spectral channels simultaneously.
Req2	The system shall image the Earth surface without obscuration between consecutive images.
Req3	The system shall provide image data at a maximum rate of 20 Mbps.
Req4	The system shall self-command and control.
Req5	The system shall have a performance (MTF, resolution or similar) better than 5 units.
Req6	The system shall have a SSD lower than 2 nm.
Req7	The system shall have a power consumption lower than 200 W.
Req8	The system shall operate at 650 km altitude and 70° inclination.
Req9	The system shall have a mass lower than 950 kg.
Req10	The system shall fit inside an envelope of 1 m.

Table 3. Example requirement dependencies

Connection	Rationale
Req1-Req3	(n) The number of channels to be imaged influences the amount of data generated and therefore the data rate required to transfer all data.
Req1-Req5	(c) Performance requirement related to the function. Therefore, its existence is only sensible if the function has to be implemented.
Req1-Req6	(c) Performance requirement related to the function. Therefore, its existence is only sensible if the function has to be implemented.
Req1-Req7	(r) The function requires power to operate.
Req2-Req3	(n) Obscuration drives the number of images to be taken per second, which influences the amount of data generated and therefore the data rate required to transfer all data.
Req2-Req7	(r) The function requires power to operate.
Req2-Req8	(p) (n) The satellite orbit speed depends on its altitude. Varying orbit speed results in different sizes of image taking, influencing therefore obscuration. For higher orbits effect is positive whereas for lower orbits effect is negative.
Req3-Req5	(n) Higher resolution requires higher data rates.
Req3-Req7	(p) Lower data rates require lower power to function and vice versa.
Req4-Req7	(r) The function requires power to operate.
Req4-Req9	(r) The function requires mass for its implementation.
Req4-Req10	(r) The function requires physical size for its implementation.
Req5-Req8	(p) (n) Optical performance is related to the distance to the focus, which is derived from the orbit altitude.
Req5-Req9	(r) Optical performance depends on size and mass of optical elements, as well as in thermo-mechanical properties such as stiffness, which influence the required mass.
Req5-Req10	(r) Optical performance depends on size of the optical elements.
Req8-Req9	(p) (n) Orbit altitude determines space environment, which influences the mass required for example for protection.
Req8-Req10	(p) (n) Orbit altitude determines space environment, which influences the size required for example for protection.

- *Temporal*: requirements are temporally related.

Kulshreshtha, Boardman, and Verma (2012) synthesize their findings in literature by enclosing all proposed dependence relations in 7 types:

- *Requires*: indicates a unidirectional dependency in which one requirement needs another one to be fulfilled.
- *Requires (loop)*: indicates a bidirectional dependency in which one or more requirements need the other ones to be fulfilled.
- *Implementation sequence*: indicates a requirement needs another one to be fulfilled before it can be fulfilled itself.
- *Value/cost*: fulfillment of a requirement affects the value of another requirement.
- *Derive*: a requirement is derived from another requirement.
- *Structure*: requirements are similar.
- *Conflict*: requirements cannot be implemented together.

However, Kulshreshtha, Boardman, and Verma (2012) identify a need to provide a classification or mechanism that is able to determine and represent dependencies among different levels of the requirement partitioning. In particular the authors come up with five categories:

- *Contractual*: transfer of data, which is equivalent to *Requires and Requires (loop)*.
- *Continuance*: continuation of flow of activities in sequence, passing message without transfer of data to help process flow.
- *Compliance*: implementation in compliance with a law/policy/rule of the company, government, or industry, which is equivalent to *Requires*.
- *Cooperation*: objective of a dependent non-functional requirement to be achieved by other requirements, which is equivalent to *Requires*.
- *Consequential*: change in modifier forcing change in other requirements, which is equivalent to *Value/Cost*.

Eben and Lindemann (2010) provide a different perspective on connectivity of requirements and focus on classifying the nodes and groups of nodes instead of their links. They propose 9 criteria to classify the single nodes and 7 criteria to classify the subsets of nodes, based on the amount and effect of their links on other nodes.

Representation of connectivity for evaluating change propagation has also been addressed by the research community. Keller et al. (2005) evaluate four techniques for visualizing change propagation through product components:

Table 4. Uncertainties affecting requirements during design phase for a space system

Uncertainty	Rationale or Example
Market size	New estimations on market size may result in adaptation of requirements.
Competitor	Introduction of competitors in the market may result in adaptation of requirements to be more competitive.
Schedule	The longer it takes to develop a system, the more probable stakeholders may change requirements.
Cost	Variation in cost may lead to modify (upgrade or waive) requirements.
Technical capability	Technical capability of the manufacturer may lead to modify (upgrade or waive) requirements.
Customer involvement	The more the customer is involved the more probable requirements will evolve.
Export	Export regulations may result in updating requirements.
Frequency allocation	Export regulations or frequency restrictions may result in updating requirements.
Mission-specific regulations	Mission-specific regulations may evolve and result in updating requirements.
Disposal	Evolution of disposal regulations may result in updating requirements.

- Design Structure Matrix (DSM): dependencies are represented in a matrix form, easing the direct linkage between different components, but its binary information provides little information on how the changes actually propagate.
- Change Risk Plot: based on the DSM concept binary information is substituted by risk information, providing a visual perception of the critical areas.
- Propagation Networks: dependencies are represented in a network form, being each component a hub or element on the network and the particular dependency a link. Links can be of different types. The graphical representation allows for making appropriate layout clustering zones according to their criticality.
- Propagation Tree: based on the Propagation Networks, the propagation tree repeats nodes for different types of links, i.e. each pair of nodes can only have one link. Although for a higher number of components the diagram gets more complex it provides a better visual representation of change propagation.

All the proposed methods can be easily transferred to evaluate propagation of requirements change, being then the requirements treated as components (nodes) and their dependencies as links.

CASE STUDY

Definitions and assumptions

Requirement connectivity is represented by a network of interconnected hubs, in which each hub represents a requirement. Each connection in the network represents a dependency between the hubs (requirements). Elements without any connection indicate orthogonal requirements, i.e., independence of requirements. Because connectivity is going to be used to evaluate conflict related to change types of connection (dependencies) are defined according to the classification proposed by Robinson, Pawloaski, and Volkov (1999), which has been shortly presented earlier. The following nomenclature is used: p (positive correlation), n (negative correlation), r (resource), c (causality).

Let's consider a space instrument with requirements as listed in Table 2. Its mission is to take images of the Earth in several spectral bands without obscuration between different data takes.

The dependency network is assessed by analyzing the dependency of each requirement with every other requirement. Dependencies in Table 3 have been determined.

For visualization purposes the Propagation Network is used in this paper to represent the dependencies between the requirements of the case study.

Values next to each hub (requirement) represent their level of connectivity, that is, how many requirements have a relation to that one. It shall be noted that connectivity accounts only those links that make an influence. For example, causal relation is only counted in one direction, as the effect requirement has not impact on the cause one. This is shown by the arrowed links. In addition, the reader shall bear in mind that connectivity in requirements represents effect on the fulfillment of a requirement and not on the implementation of the component to which the requirement is allocated.

The resulting values can then be used as a prioritization factor for decision-making. In essence connectivity indicates how many requirements may be affected when one of them faces change.

Application

Salado, Nilchiani, and Efatmaneshnikh 2012 identify and categorize major uncertainties that may affect requirements during the design phase of a space system once a triggering event occurs, which are listed in Table 4.

In the following the connectivity diagram presented in Figure 2 is used to assess how requirements are affected when uncertain events occur. Several cases are studied, one per each defined uncertainty.

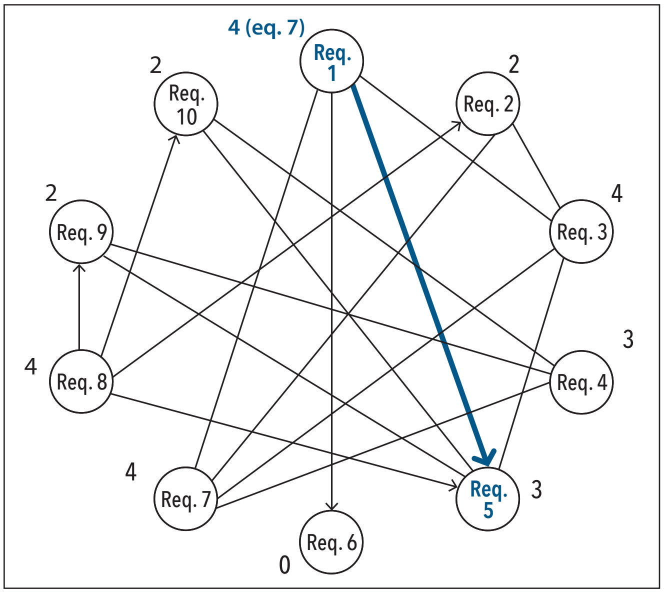


Figure 2. Visual representation of requirement dependencies including effects of causal relations

Due to the scope of the system under study, a space-based Earth observation instrument, the following uncertainties are considered not applicable, as they influence other parts of the system or the higher level system, but not the instrument itself: frequency allocation (the instrument is not in charge of communicating to Earth or an external system), mission-specific regulations (they usually have to do with the Concept of Operations of the higher level system), and disposal (managed by Concept of Operations and affecting the platform, not the instrument).

Note on color code for the network diagrams that follow: requirements marked in red are the ones changing. Requirements marked in yellow are the ones being affected. Red lines and arrows represent probable negative impacts (more stringent requirement). Green lines and arrows represent additional margin to fulfill a requirement.

Case 1 – Market size. New estimations of the market size show a less optimistic prediction than initially planned. As a result, investors decide to change the satellite platform in order to reduce upfront investment. It results in lower resources for the instrument, namely lower data rate (Req. 3), lower available power (Req. 7), and lower mass capability (Req. 9). The effect of each requirement change in the rest of the requirements is shown in Figure 3.

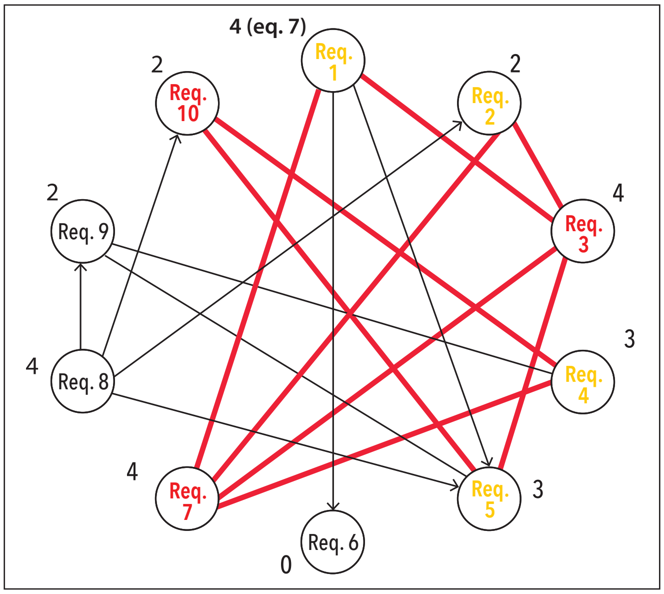


Figure 3. Uncertain triggering event on requirements for Case 1

Power availability (Req. 7) constraints the amount of functions that can operate simultaneously (Req. 1, Req. 2 and Req. 4) and in some cases may limit performance (Req. 3). In optical instruments mass restrictions (Req. 10) constraint the size of optical elements, which influences optical performance (Req. 5). The dependency analysis shows that change in resources conditions fulfillment of functionality and performance. A joint connectivity of the resource requirements of 9 affects 5 requirements through 9 channels.

Case 2 – Competitor. During the course of system development a competitor has put into service a space system at the same resolution the instrument is being developed to. Therefore investors decide to upgrade the design and increase the required resolution (Req. 5). The effect of modifying such requirement is shown in Figure 4.

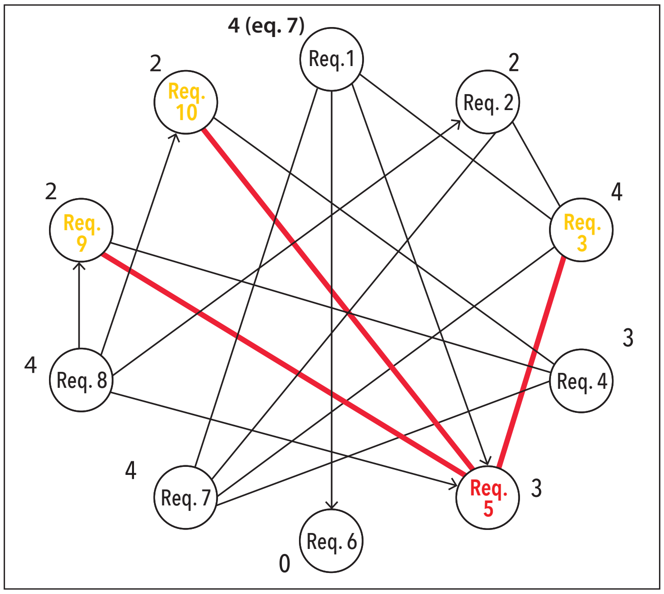


Figure 4. Uncertain triggering event on requirements for Case 2

DECEMBER 2024
VOLUME 27/ ISSUE 6

Fulfillment of more stringent optical requirements (Req. 5) may require bigger and heavier components. Therefore mass (Req. 9) and envelop (Req. 10) requirements may not be fulfilled. The dependency analysis shows that change in performance conditions fulfillment of resources.

A connectivity of 3 affects 3 requirements through 3 channels.

Case 3 – Schedule. During the course of system development scientists have come up with a new processing algorithm that brings different benefits to the analysis of Earth images. This algorithm requires satellites to take overlapping images during passes. In order to take benefit from this use investors have decided to change the obscuration requirement (Req. 2) and instead require the instrument to take consecutive images

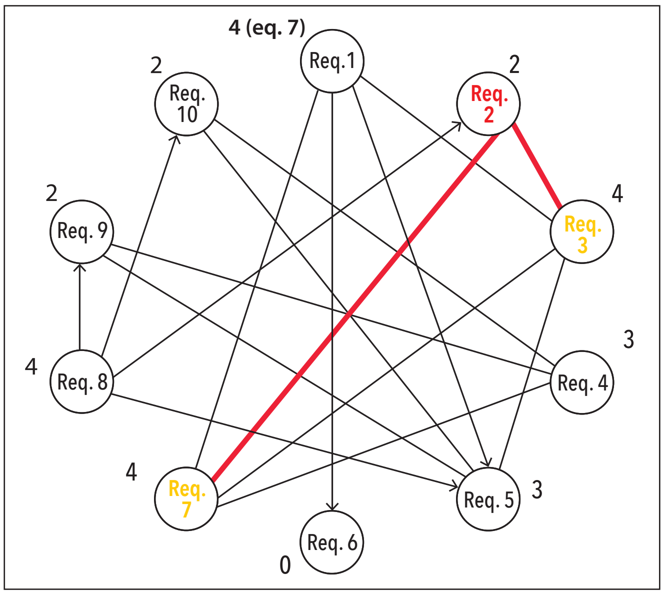


Figure 5. Uncertain triggering event on requirements for Case 3

ensuring at least 60% overlap. The effect of modifying such requirement is shown in Figure 5.

Increasing amount of images per second (Req. 2) may require higher power (Req. 7). Since significantly more data are produced, fulfillment of maximum data rate (Req. 3) is put into question. The dependency analysis shows that change in functionality conditions fulfillment of performance and resources. A connectivity of 2 affects 2 requirements through 2 channels.

Case 4 – Cost. Due to severe cost overrun during development of the instrument system the principal contractor has decided to move command and control capabilities (Req. 4) back to the satellite platform in order to limit equipment investment at instrument side (at the risk of having a more complex verification and validation). The effect of modifying such requirement is shown in Figure 6.

Removing functionality from the system (Req. 4) may result in removing system components, which frees the availability of power, mass and volume (Req. 7, Req. 9, and Req. 10). The dependency analysis shows that change in functionality conditions fulfillment of resources. A connectivity of 3 affects 3 requirements through 3 channels.

Case 5 – Technical capability. During instrument development the instrument manufacturer realizes it cannot achieve the

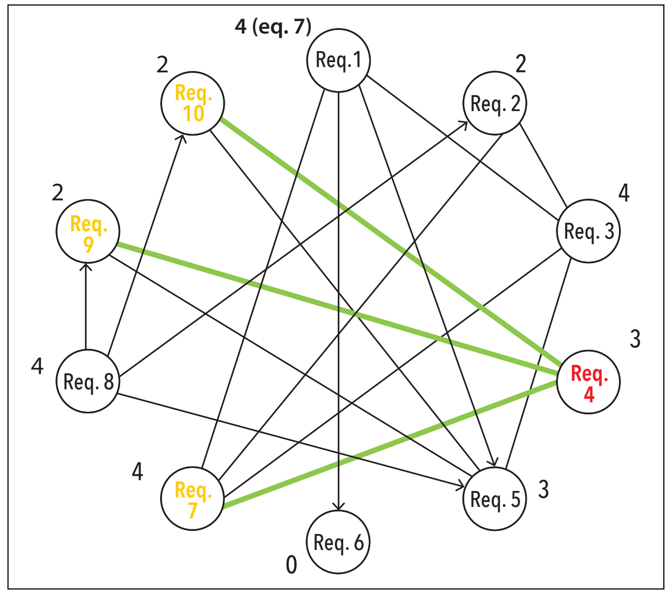


Figure 6. Uncertain triggering event on requirements for Case 4

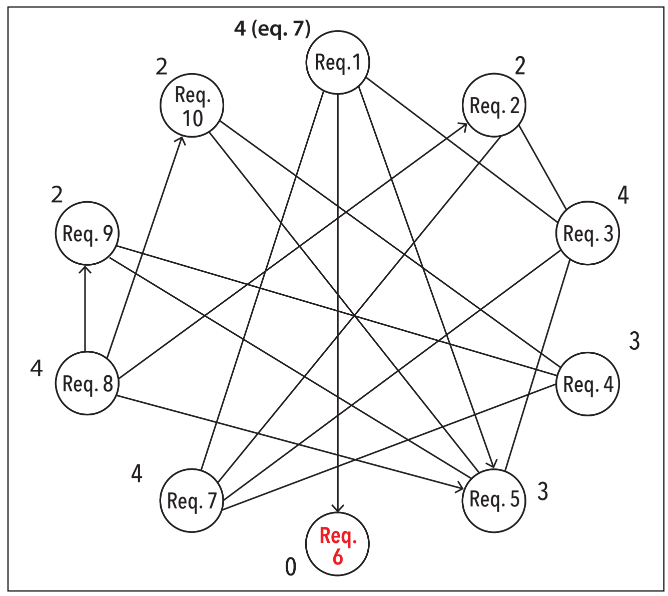


Figure 7. Uncertain triggering event on requirements for Case 5

required SSD (Req. 6). As a consequence they issue a Request For Deviation (RFD) to the customer, which is approved. The effect of modifying such requirement is shown in Figure 7.

Since SSD is associated to the physical properties of a material, its fulfillment does not have any impact on other requirements. A connectivity of 0 does not affect any requirement.

Case 6 – Customer involvement. In this particular project customer is significantly involved in all levels of the system development. After reviewing the operational concept of the instrument and evaluating test results of some lower level components, customer has decided to add a new requirement in order to acquire health status data of the instrument for post-processing when operational. The following requirement is added: **Req. 11.** *The system shall provide the housekeeping data according to REF1.* The effect of adding such requirement is shown in Figure 8.

Adding housekeeping functionality to the system (Req. 11) may

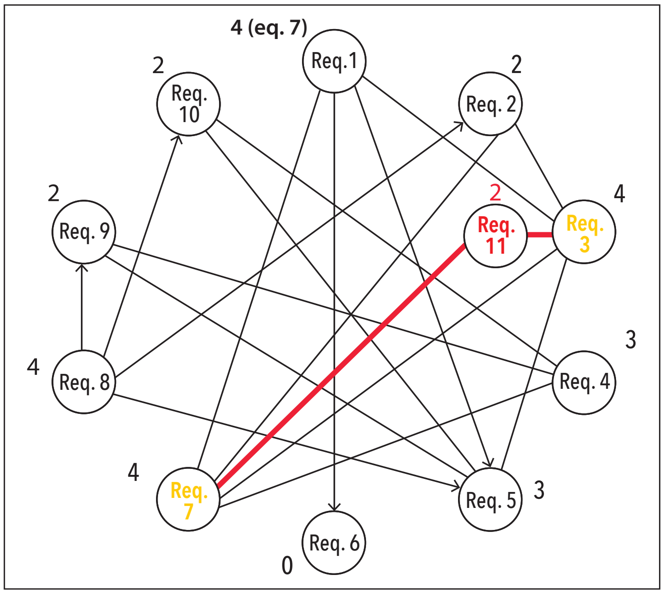


Figure 8. Uncertain triggering event on requirements for Case 6

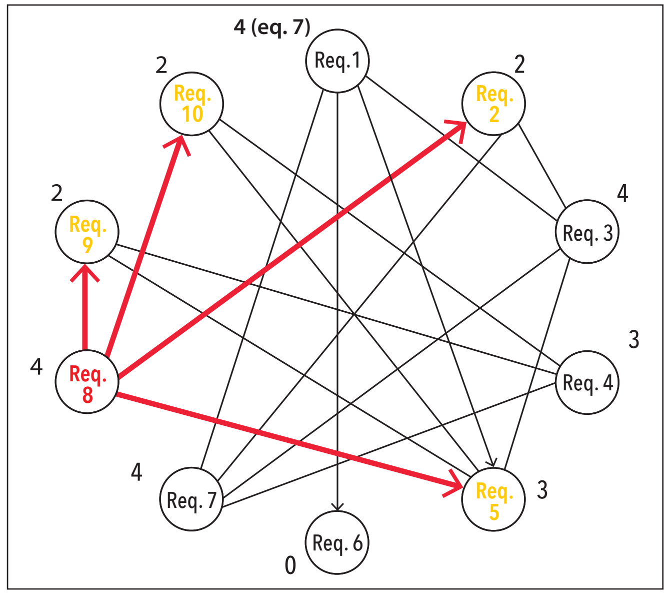


Figure 9. Uncertain triggering event on requirements for Case 7

require additional power (Req. 7). Since additional data are generated, fulfillment of data rates may be questionable (Req. 3). The dependency analysis shows that change in functionality conditions fulfillment of resources and performance. A connectivity of 2 affects 2 requirements through 2 channels.

Case 7 – Export. During instrument development USA export control regulations changed, impacting the legality to use the pre-selected launcher. In order to avoid incrementing the complexity of in-orbit maneuvers investors have decided to use a different orbit to park their satellite (Req. 8). The effect of modifying such requirement is shown in Figure 9.

Changing the orbit in which the instrument operates (Req. 8) directly influences optical performance (Req. 5), as the optical focus of the system changes. In addition, the different orbit position may risk the fulfillment of the obscuration requirement (Req. 2), as system may be unable to record images at certain rate. Furthermore, a different orbit may present different radiation environment, resulting in the need of additional mass and size (Req. 9 and Req. 10) to incorporate the appropriate level of radiation hardening. The dependency analysis shows that change in interaction conditions fulfillment of functionality, performance, and resources.

A connectivity of 4 affects 4 requirements through 4 channels.

Summary of results

The results of the different cases are summarized in Table 5.

CONCLUSIONS

The present research shows that requirement connectivity can be effectively used to assess the impacts of uncertain events on system requirements during the entire life-cycle of a project. The impact of requirement changes can be illustrated with help of the presented approach; it thus facilitates impact analysis. The following patterns and conclusions can be recognized in the case-study described in the paper:

- Connectivity of a causal requirement is related to the amount of affecting channels.
- Number of affected requirements is lower or equal to the connectivity of a requirement.
- Apparently, there is certain correlation between the type of the causal requirement and the type of the affected requirements.
- There is no apparent correlation between the type of uncertainty and the amount of affected requirements.

The present research has set the basis for future research that would overcome the limitations of the results presented in this paper:

- Apply the proposed connectivity determination method to higher amount of requirements to improve pattern recognition and include interaction requirements.

Table 5. Evaluation case analyses

Case	Causal requirement type	Dependency impacts			Affected requirement type			
		C	R	Ch	F	P	R	I
1	Resource	9	5	9	X	X		
2	Performance	3	3	3			X	
3	Function	2	2	2		X	X	
4	Function	3	3	3			X	
5	Performance	0	0	0				
6	Function	2	2	2		X	X	
7	Interaction	4	4	4	X	X	X	

* The following definitions are used in the table:
- C: connectivity of the causal requirement.
- R: number of affected requirements.
- Ch: number of affecting channels.
- F: functionality; P: performance; R: resources; I: interaction.

- Apply the proposed connectivity determination method to higher amount of cases to improve pattern recognition.
- Incorporate structural definitions into pattern recognition.
- Incorporate advances on network theory to develop efficient computation algorithms.
- Formally complete the model with mathematical representations, metrics, and definitions.
- Develop techniques to ensure completeness of the dependency identification activity. ■

REFERENCES

- Berander, P., and A. Andrews. 2005. "Engineering and managing software requirements: requirements prioritization." *Springer Journal*, pp.70–91.
- Bischof, A., and L. Blessing. 2007. "Design for flexibility: making provisions for requirement changes." International Conference on Engineering Design ICED07, 28 – 31 August, Cite des Sciences et de l'Industrie, Paris, FR.
- Bonjour, E., G. Harmel, P. Micaelli, and M. Dulmet, M. 2010. "Simulating change propagation between product architecture and development organization." *International Journal of Mass Customisation* 3, 3: 288-310.
- Carlshamere, P., K. Sandahl, M. Lindvall, B. Regnell, and N. D. Band Dag. 2001. "An industrial survey of requirements interdependencies in software product release planning." *RE '01 Proceedings of the Fifth IEEE International Symposium on Requirements Engineering.*
- Clarkson, P. J., C. Simons, and C. Eckert. 2001. "Predicting change propagation in complex design." *Proceedings of DETC'01, ASME 2001 Design Engineering Technical Conference and Computers and Information in Engineering Conference*, Pittsburgh, US-PA, 9-12 September.
- Eben, K. G. M., and U. Lindemann. 2010. "Structural analysis of requirements – Interpretation of structural criterions." *12th International Dependency and Structure Modelling Conference DSM'10*, 22 – 23 July, Cambridge, GB.
- Firesmith, D. 2004. "Prioritizing Requirements." *Journal of Object Technology* Vol. 3, No. 8, September-October.
- Giffin, M., O. de Weck, G. Bounova, R. Keller, C. Eckert, and P. J. Clarkson. 2009. "Change Propagation Analysis in Complex Technical Systems." *Journal of Mechanical Design*, August, Vol. 131 / 081001-1.
- Hassan, A. E., and R. C. Holt. 2004. "Predicting Change Propagation in Software Systems." *Proceedings of the 20th IEEE International Conference on Software Maintenance*, Pages 284-293, IEEE Computer Society, Washington, US-DC.
- Hull, E., K. Jackson, and J. Dick. 2005. *Requirements engineering.* Second edition, Springer, ISBN:
- 1-85233-879-2.
- Karlsson, J. 1996. "Software requirements prioritizing." *Proceedings of the Second International Conference on Requirements Engineering (ICRE'96)*, IEEE Computer Society, Colorado Springs, US-CO, 15–18 April, Los Alamitos, US-CA, pp.110–116.
- Keller, R., C. M. Eckert, and P. J. Clarkson. 2006. "Heuristics for change prediction." *International Design Conference DESIGN 2006*, Dubrovnik, Croatia, 15-18 May.
- Keller, R., T. Edger, C. M. Eckert, and P. J. Clarkson. 2005. "Visualising change propagation." *International Conference on Engineering Design ICED 05*, Melbourne, AU, 15-18 August.
- Koh, E. C. Y, and P. J. Clarkson. 2009. "A modeling method to manage change propagation." *International Conference on Engineering Design ICED'09*, 24-27 August, Stanford University, Stanford, US-CA.
- Kulshreshtha, V., J. Boardman, and D. Verma, D. 2012. "The emergence of requirements networks: the case for requirements inter-dependencies." *Int. J. Computer Applications in Technology*, Vol. 45, No. 1, 2012 (to be published).
- Mead, N. R. 2006. *Requirements prioritization introductio*n. Software Engineering Institute, Carnegie Mellon University, Pittsburgh, US-PA.
- Oduncuoglu, A., and V. Thomson. 2011. "Evaluating the risk of change propagation." *International Conference on Engineering Design ICED'11*, 15-18 August, Technical University of Denmark.
- Peterson, C., R. K. Paasch, P. Ge, and Dietterich. 2007. "Product innovation for interdisciplinary design under changing requirements." *International Conference on Engineering Design ICED07*, 28 – 31 August, Cite des Sciences et de l'Industrie, Paris, FR.
- Pohl, K. 1996. *Process-Centered Requirements Engineering.* John Wiley & Sons, Inc. New York, US-NY, ISBN: 0863801935.
- Robinson, W. N., S. D. Pawloaski, and V. Volkov. 1999. "Requirement interaction management." http://www.cis.gsu.edu/~wrobinso/papers/GSUCIS99-7.PDF. Department of Computer Information Systems, Georgia State University, Atlanta, US-GA, wrobinson@gsu.edu (This online publication was first accessed around December 2008).
- Salado, A., R. Nilchiani, and M. Efatmaneshnikh. 2012. "Taxonomy and Categorization of Uncertainties in Space Systems with an Application to the Measurement of the Value of Adaptability." *AIAA Space Conference*, Pasadena, US-CA.
- Sudin, M. N., and S. Ahmed-Kristensen. 2011. "Change in requirements during the design process." *International Conference on Engineering Design ICED11*, 15 – 18 August, Technical University of Denmark.
- Zhan, W., H. Mei, and H. Zhao, H. 2005. "A feature-oriented approach to modeling requirements dependencies." *Proceedings of 2005 13th International Conference on Requirement Engineering (RE'05)* IEEE.

ABOUT THE AUTHORS

[Editor: Author biographies were current when the paper was initially published in 2013.]

Alejandro Salado is a doctoral candidate with the School of Systems and Enterprises at the Stevens Institute of Technology, for which he was awarded the International Fulbright Science and Technology Award. His research interests include the generation and prioritization of requirements, the design of elegant systems, and the development of affordable space systems. Experienced in the development of manned and unmanned space systems, he is a senior systems engineer at Kayser-Threde GmbH (an OHB company) in Germany, currently acting as verification manager for the EnMAP satellite, actively contributing to the generation of the company's systems engineering capability, and leading an initiative to transform it towards model-centricity. Alejandro holds a MSc in electrical engineering from Polytechnic University of Valencia, a MSc in electronics engineering and a MSc in project management from Polytechnic University of Catalonia, and the SPACETECH postgraduate Master in space systems engineering from the Delft University of Technology.

> continued on page 60

Applying Bayesian Networks to TRL Assessments – Innovation in Systems Engineering

Marc F. Austin, Virginia Ahalt, Erin Doolittle, Cheyne Homberger, George A. Polacek, and **Donal M. York**

ABSTRACT
Currently, technology readiness assessments (TRAs) are used in determining the maturity of the critical technology elements (CTEs) of a system as it moves forward in the system development life cycle. The TRA method uses technology readiness levels (TRLs) as the decision metric. TRL values are assessed and determined by subject matter experts (SMEs). Since expert evaluators often differ in their judgment when scoring a system element against the TRL scale criteria, this paper argues for the use of a Bayesian network model to provide a mathematical method to consistently combine and validate the judgment of these SMEs and increase the confidence in the determination of the readiness of system components and their technologies.

INTRODUCTION

Bayesian probability provides a framework for building and refining models which incorporate uncertainty. A Bayesian network, or Bayes net, is a graphical representation of a multi-dimensional probability distribution, in which a variety of indicators may be dependent on a complex network of observable and hidden variables. Bayes nets are well suited for translating complex relationships of dependencies into intuitive and mathematical models and perform well even in the face of missing or inconsistent data.

In this case, our challenge is the decision-making process that assigns technology readiness level (TRL) values to system technologies or critical technology elements (CTEs) in the DoD's technology readiness assessment (TRA) process. In performing a TRA, assessors consider a variety of different and often subjective attributes of a system in order to make a final determination, which is as consistent as possible. The Bayes net effectively models this situation: it is able to incorporate a set of complex, possibly incomplete, and highly interrelated attributes and, through

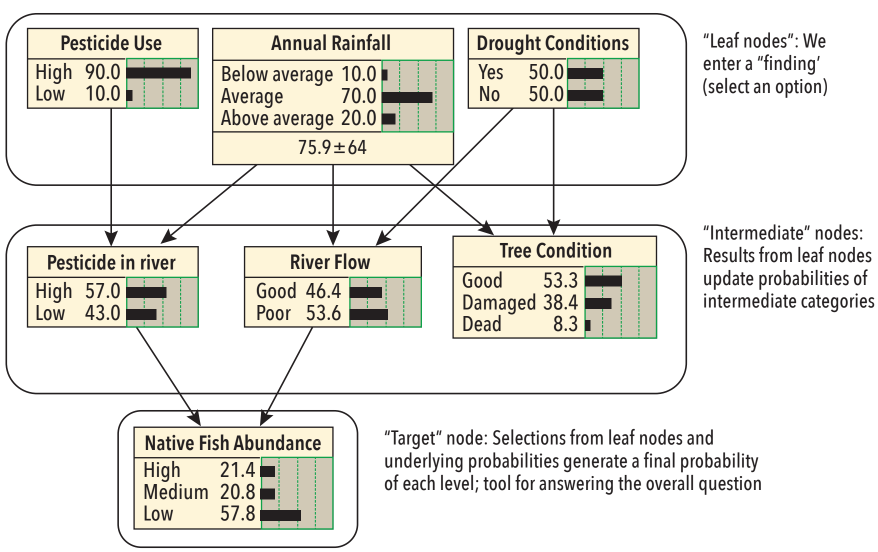


Figure 1. An example Bayesian belief network – predicting native fish abundance

Table 1. Decision criteria for assessing technology readiness level (TRL[1])

TRL	Definition	Description	Supporting Information
5	Component and/or breadboard validation in a relevant environment.	Fidelity of breadboard technology increases significantly. The basic technological components are integrated with reasonably realistic supporting elements to they can be tested in a simulated environment. Examples include "high-fidelity"laboratory integration of components.	Results from testing laboratory breadboard system are integrated with other supporting elements in a simulated operational environment. How does the "relevant environment" differ from the expected operational environment? How do the test results compare with expectations? What problems, if any, were encountered? Was the breadboard system refined to more nearly match the expected system goals?
6	System/subsystem model or prototype demonstration in a relevant environment.	Representative model or prototype system, which is well beyond that of TRL 5, is tested in a relevant environment. Represents a major step up in a technology's demonstrated readiness. Examples include testing a prototype in a high-fidelity laboratory environment or in a simulated operational environment.	Results from laboratory testing of a prototype system that is near the desired configuration in terms of performance, weight, and volume. How did the test environment differ from the operational environment? Who performed the tests? How did the test compare with expectations? What problems, if any, were encountered? What are/were the plans, options, or actions to resolve problems before moving to the next level?
7	System prototype demonstration in an operational environment.	Prototype near or at planned operational system. Represents a major step up from TRL 6 by requiring demonstration of an actual system prototype in an operational environment (e.g., in an aircraft, in a vehicle, or in space).	Results from testing a prototype system in an operational environment. Who performed the tests? How did the test compare with expectations? What problems, if any, were encountered? What are/were the plans, options, or actions to resolve problems before moving to the next level?

1. Developed by NASA and recommended by the Defense Acquisition Guidebook.

the laws of probability, produce a consistent and mathematically rigorous recommendation. The model is constructed through gathering evidence and eliciting expert opinion, which are all incorporated, along with any uncertainty, in the final product. Each individual indicator is represented as a node in the network, with links representing dependencies between the nodes. Bayes theorem governs the relationships between the connected nodes. Figure 1 illustrates an example Bayesian belief network.

WHY USE BAYESIAN NETWORKS FOR TRLS?

The Bayes net provides an effective framework for testing the most likely outcome of future events or scenarios and finding their likely cause(s). It combines both subjective expert opinions with available quantitative information/data providing informed decision-making without requiring complete knowledge of the problem. The problem domain experts take ownership because their input is vital. As new knowledge is acquired, the modular design of the Bayes net easily accommodates additional information.

The TRL is a systematic metric/measurement to assess the maturity of a particular technology and to allow consistent comparisons of maturity between different types of technologies. The TRL was initially pioneered by J. C. Mankins [6] at the National Aeronautics and Space Administration (NASA) Goddard Space Flight Center in the 1980s as a method to assess the readiness and risk of space technology. Over time, NASA continued to use readiness levels as part of an overall risk assessment process and as a means for comparison of maturity between various technologies. NASA incorporated the TRL methodology into the NASA Management Instruction 7100 as a systematic approach to the technology planning process. The DoD, along with several other organizations, later adopted this metric and tailored its definitions to meet their needs. TRL values range from 1 to 9. A definition, description, and decision criteria for TRL values 5, 6, and 7 are provided in Table 1. The technology readiness assessment (TRA) method uses TRLs and is a Department of Defense (DoD) directive performed across the DoD [7]. A TRL of 6 is a particularly critical milestone in the US DoD systems development life cycle as it is required to enter full scale development.

In the technology readiness assessment, critical technology elements or CTEs are selected from among the elements/components of the development system. The TRL of each of these CTEs is assessed by subject matter experts (SMEs). Although scoring a technology in conjunction with the 1 to 9 TRL scale is based on satisfying certain requirements and providing the accompanying evidence, expert evaluators may often differ in their judgment. Use of the Bayesian network and its resulting probability distributions help to validate and establish a level of confidence in the judgment of these experts. The TRL Bayesian network model yields a distribution of TRL values and therefore represents or is suitable for analyzing a variety of scenarios. A range of TRL values rather than a single number provides the analyst with a level of confidence and a better perspective of where the risk lies.

CONSTRUCTING THE BAYESIAN NETWORK

Figure 2 illustrates the process steps followed in constructing the TRL Bayesian network. The question of interest is "*What is the technology readiness level of the technology element?*" First, identify everything that contributes to determining a TRL, i.e., the set of variables or nodes. Is there a natural ordering of these nodes? Can the nodes be treated as having binary states? Next, what is the dependency structure? Ascertain how to preserve an acyclic requirement and how to significantly reduce complexity. Finally, what is the conditional probability table (CPT) for each node?

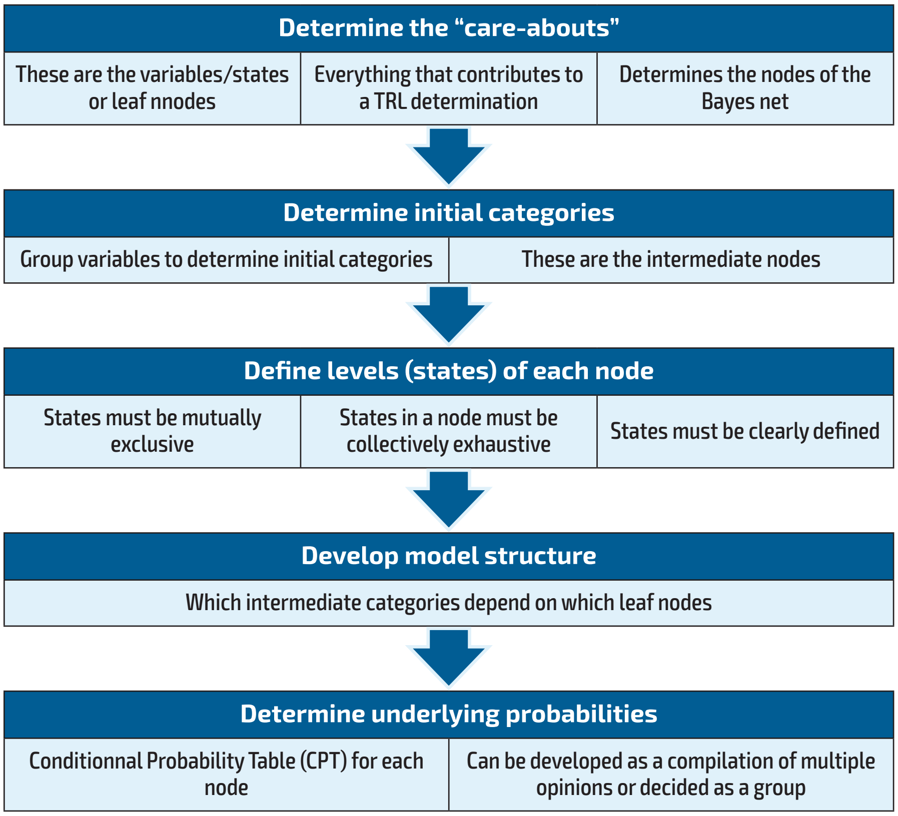


Figure 2. Constructing a Bayesian Network

There are certain requirements for a good variable. The values (or states) must be mutually exclusive. In other words, two or more states cannot be true at any single time. Also, states in a variable must be collectively exhaustive. One of the states must be true. Lastly, states must be unambiguously defined. Ambiguous states, such as other, should be avoided. Use a clarity test for each variable. Clarity means that a person looking at the variable knows without interpretation or assumption what the variable means and what its value is.

Figure 3 shows the complete set of TRL care-abouts or variables that are eventually reduced to the leaf nodes in the model. These variables resulted from a series of brainstorming meetings of the SMEs to list everything they thought contributed to or might potentially impact the decision-making process determining a TRL. While the conventional TRL tables define the TRL at each level, the purpose of the care-abouts is to capture a much more comprehensive and detailed set of attributes considered by the SME in their decision-making process. Along the way, numerous variables were combined, deduped or determined not to influence the TRL decision and were

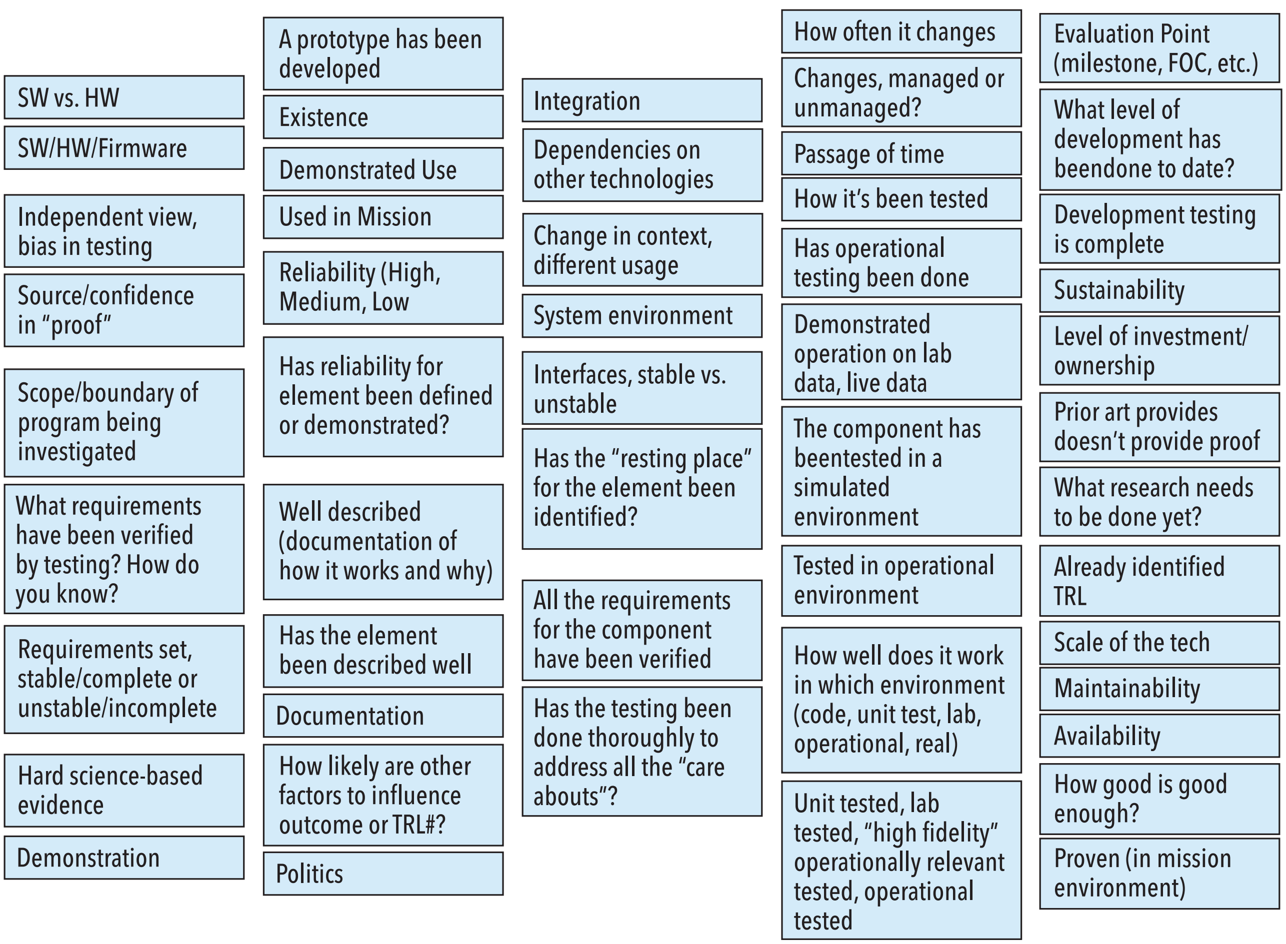


Figure 3. The initial set of variables or care-abouts from the TRL brainstorming meetings

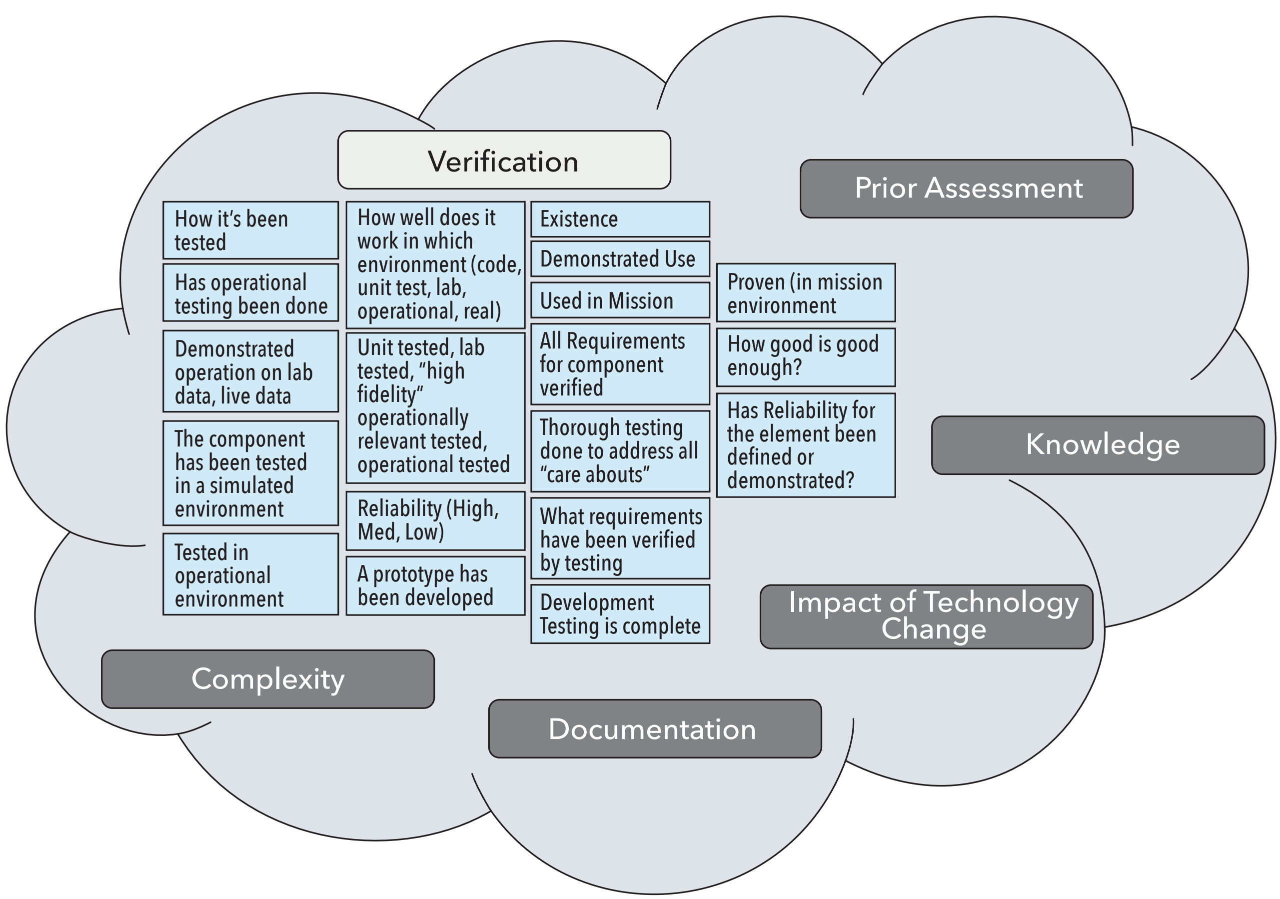


Figure 4. A snapshot of the Bayesian network categories and specifically the sub-categories for verification

eliminated. During this process, focus is placed on the as is condition and factors that attempted to predict the future were intentionally removed from the initial set. Examples of the latter would be Politics and Sustainability.

The variables shown in Figure 3 are then grouped into a set of initial categories which determine the intermediate nodes in the Bayesian belief network. For example, from the factors identified, a category that manifested itself early in the analysis was verification. Figure 4 shows the initial set of care-abouts that were collected into this the verification category. As will be seen later the verification category was ultimately reduced to two sub-factors, test environments and level of testing passed.

Once the initial categories are determined, the levels (states) of each node are clearly defined. States must be mutually exclusive. Figure 5 shows the main category *knowledge* and its three sub-nodes or sub-categories: *research, proof-of-concept*, and *prior usage*. Figure 4 reflects the progression of the development of the model structure as it portrays which intermediate categories depend on which leaf nodes.

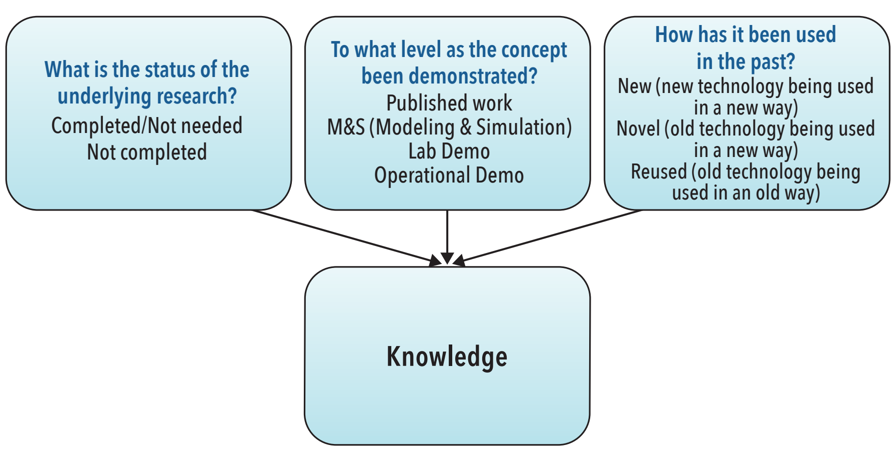


Figure 5. The knowledge category and its sub-categories for the Bayesian network model

The final step in the construction of the Bayesian network is to determine the underlying probabilities and construct a conditional probability table (CPT) for each node. Since no priors or previous underlying distributions existed in the case of the TRA process, the conditional probabilities for the nodes were elicited and compiled from Subject Matter Experts. Table 2 provides some example entries in the CPT for the *impact of technology change* node. The interpretation of the table says given the conditions in the first three columns, how *likely* is it that the state of the node (in this example, the *impact of technology change*) is classified as High? Moderate? or Low?

With a basis in Bayes theorem and the laws of probability, the overall model shown in Figure 6 is constructed by linking the individual CPTs together in accordance with the structure of the model, culminating in the final TRL value. Each individual variable then has predictive power over the final result, but the influence is interdependent on the values of each of the other nodes. That is, a change to a single finding may have different results on the final TRL

Table 2. Sample conditional probability table (CPT) for the impact of technology change node

Control of Technology Change	Frequency of Technology Change	Magnitude of Technology Change	Explanation	High	Moderate	Low
Managed	Frequent	Large	If **managed** technology change **frequently** occurs and the magnitude of that change is **large** how likely is the impact of that technology change to be: **High? Moderate? Low?**	33	34	33
Managed	Never	Small	If **managed** technology change **never occurs** and the magnitude of that change is small how likely is the impact of that technology change to be: **High? Moderate? Low?**	0	0	100
Unmanaged	Infrequent/ Seldom	Small	If **unmanaged** technology change **infrequently/seldom** occurs and the magnitude of that change is **small** how likely is the impact of that technology change to be: **High? Moderate? Low?**	60	30	10

assessment depending on the states of the other nodes. For example, the quality of documentation may have a much stronger impact on the eventual assessment for an immature system than one which has already undergone testing.

MODEL CASE STUDIES

Technology readiness assessments (TRAs) from existing or prior programs were sought in order to validate the Bayesian network model. The questionnaire shown in Table 3 reflects the model structure shown in Figure 6. Nodes are reflected as a *category* and a sub-node as a *subcategory*. The responses to the individual questions of the questionnaire determine the values for each of the model's leaf nodes (see Figure 6) which are then combined using Bayes theorem and the laws of probability. The questionnaire in Table 3 below reflects a completed evaluation for a project whose TRL assessment was a 3. No specific details are provided for any of the model case studies as the projects and systems involved were proprietary in nature. It is important to keep in mind that the model is not intended to select a specific TRL value or to replace the expert or judgment of the expert. The model serves as a decision aid, not a decision maker. One should not compare the model against the systems engineering decision

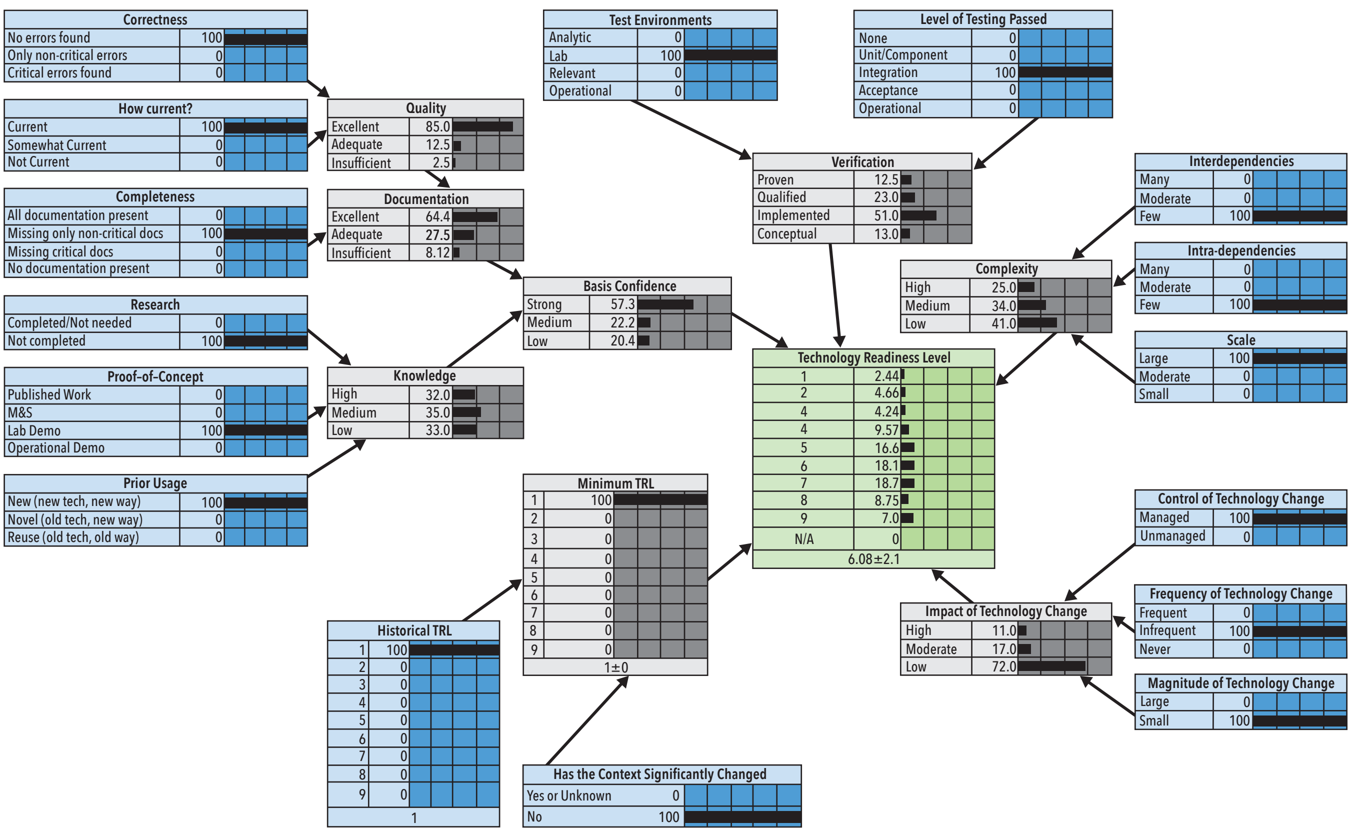


Figure 6. The final Bayesian network model

Questionnaire:

Table 3. Bayesian Network Project Questionnaire

Category	Subcategory	Question	Possible Answers (Select/highlight one)
Documentation	Quality	Correctness: How **accurate** is the documentation?	■ No Errors Found ☐ Only non-critical errors found ☐ Critical Errors Found
		How current: How **current** is the documentation?	☐ Current ■ Somewhat current ☐ Not current
	Completeness	How **complete** is the documentation?	☐ All Documentation Present ■ Missing only non-critical documents ☐ Missing critical documents ☐ No documentation present
Documentation *includes e.g., acquisition documents, architecture products, engineering specs, test plans, and general references.* ***Critical errors*** *are those which cause a misunderstanding of the facts and significantly impact the outcome.* ***A critical document*** *is any document that contains data elements essential to understanding the technology under evaluation.*			
Knowledge	Research	What is the status of the underlying research?	☐ Completed/Not needed ■ Not completed
	Proof-of-Concept	To what level has the concept been demonstrated?	☐ Published work ■ M&S (Modeling and Simulation) ☐ Lab Demo ☐ Operational Demo
	Prior Usage	How has it been used in the past?	■ New (new technology being used in a new way) ☐ Novel (old technology being used in a new way) ☐ Reused (old technology being used in an old way)
Prior Assessment	Historical TRL	Was it previously assessed at a certain TRL level? If so, what level?	■ 2 (Insert previous TRL 1-9 here, or leave blank if not previously assessed)
	Context Change	Has the context significantly changed (from prior assessment)?	■ Yes or Unknown (or, not previously assessed) ☐ No
Context change *must be linked to a prior TRL assessment. If no prior assessment occurred, select "Yes or Unknown."*			
Impact of Technology Change	Magnitude	What is the magnitude of technology change?	■ Large ☐ Small
	Frequency	What is the frequency of technology change?	☐ Frequent ■ Infrequent/Seldom ☐ Never
	Control	What is the control of technology change? (How well is change controlled?)	■ Managed ☐ Unmanaged

Table 3. Bayesian Network Project Questionnaire (continued)

Category	Subcategory	Question	Possible Answers (Select/highlight one)
Complexity	Scale	What is the scale?	■ Large ☐ Moderate ☐ Small
	Intra-dependencies	How many intradependencies?	■ Many ☐ Moderate ☐ Few
	Inter-dependencies	How many interdependencies?	■ Many ☐ Moderate ☐ Few
Scale *refers to, for example, the scope, magnitude, quantity, or breadth of the technology within the system.* ***Intra-dependencies*** *are within the technology.* ***Interdependencies*** are between the technologies.			
Verification	Test Environments	What is the environment in which the testing was conducted?	☐ Analytic ☐ Lab ■ Relevant ☐ Operational
	Level of Testing Passed	What level of testing has been passed?	■ None ☐ Unit/Component Testing ☐ Integration testing ☐ Acceptance testing ☐ Operational testing
Level of testing *refers to the highest level of testing that has been fully completed and successfully passed, with accompanying evidence.* *Testing need not be comprehensive to be completed.*			

process but rather view the model as reinforcing that process. The model can be used beforehand to help frame the assessment process or afterwards to support validation of the process and to perform risk evaluation. The following questionnaire instructions were given to the survey respondents.

> ***Questionnaire Instructions:*** *We are exploring the use of a Bayesian network approach to enhance the current method of performing technology readiness assessments (TRAs). The model uses the responses to the series of questions shown in the table that follows to ultimately generate a probability distribution of the TRL levels. Your feedback is important as it will be used to validate our model and assist us in determining the way forward.*

Using the program response as input data, the Bayes net model produces the resulting TRL probability distribution shown in Figure 7. As one can see the Bayes net model predicts about a 38% probability that the TRL = 2 and about a 30% probability that the TRL = 3. As was stated initially, the project had been assessed at a TRL = 3. One can also describe the results in terms of cumulative probability. In other words, there is approximately a 91% probability that the TRL is equal to or less than 3.

Another case study was provided where the system was assessed at a TRL of 7. Inputting the data from responses to the model questionnaire yielded the TRL probability distribution shown in Figure 8. It can be seen that the model indicates that there is approximately a 95% probability that the TRL is greater than a 7. In both these case studies, evidence from the probabilities of the states of the nodes is compared with the TRL assessment done by subject matter experts or expert engineers using their judgment. These experts are the decision makers who decide the TRL value. In the first case study the model indicates that the TRL is about 8% more likely equal to a value of 2 compared to the TRL of 3 assessed by the SMEs. For this first case the model prediction aligns closely with the judgment of the experts. In the second system, the model shows a TRL assessed at two levels higher than what the SMEs concluded. Here the judgment of the experts is

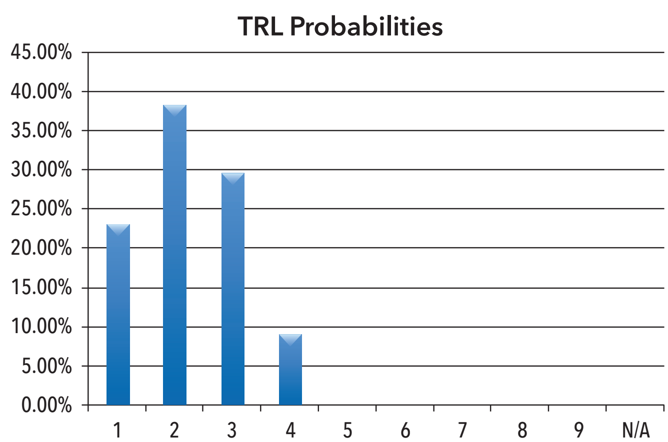


Figure 7. Bayesian network model results for a case study where the assessed TRL = 3

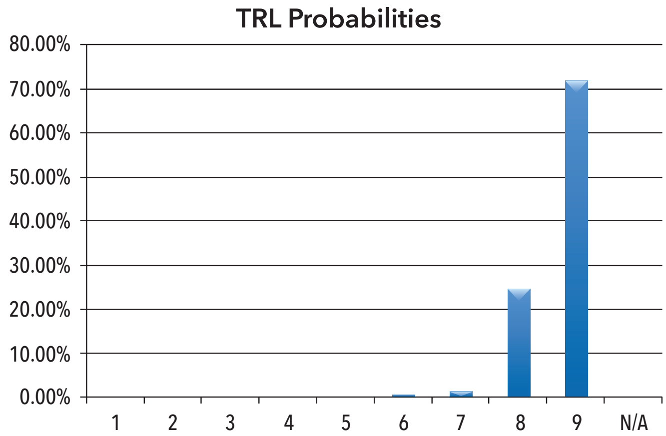


Figure 8. Bayesian network model results for a case study where the assessed TRL = 6

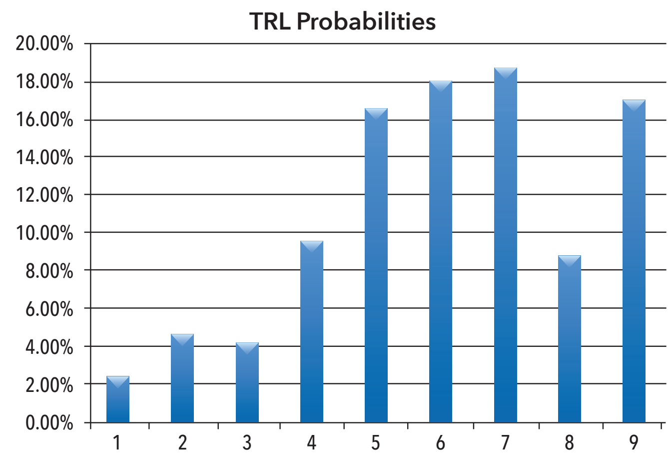


Figure 9. Bayesian network model results for a case study where the assessed TRL = 7

demonstrated to be more conservative than the model. The authors were not in a position to go back to the experts to discuss the differences with them and attempt to determine why their assessment was more risk adverse. The results from a third case study are shown in Figure 9. Here the results do not reflect a clear peak or plateau but rather a range of values. The model conservatively predicts a distribution of TRL values around the assessed TRL value of 6. Such is the case in the absence of a clear winner or with incomplete or inconsistent data, for example when the judgments of a group of experts vary.

In summary, the use of the Bayesian network model provides the systems engineer with a level of confidence in the judgments made by the SMEs in assigning a TRL. In addition to the three case studies illustrated here, the model is being used with other systems and validation work is ongoing.

FUTURE WORK

Work is currently being done to construct a Bayesian network model for integration readiness levels (IRLs). The IRL is a metric to measure the integration maturity between two or more components. IRLs, in conjunction with TRLs, form the basis for the system readiness level (SRL), a systems level metric generated from the systems readiness assessment (SRA) process [9]. The IRL values range from 0 to 9. The original IRL scale definitions, as proposed by Sauser [8], have been modified to be consistent with the foundation of the TRL scale and to reflect more closely our development model. IRLs represent the systematic analysis of the interactions between various components and provide a consistent comparison of the maturity between integration points. IRLs provide a means to reduce the risk involved in maturing and integrating components into a system. Similar to the TRL, the use of a Bayesian network model for IRLs will provide a mathematical method to consistently combine and validate the judgment of experts in the determination of the integration readiness of system components. ■

REFERENCES

- [1] Pearl, J. 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Mateo, US-CA: Morgan Kaufmann.
- [2] Jensen, F. V. 2001. *Bayesian Networks and Decision Graphs*. Springer.
- [3] Friedman, N., D. Geiger, and M. Goldszmidt. 1997. "Bayesian network classifiers." *Machine Learning* 29: 131–163.
- [4] Box, G. E. P., and G. C. Tiao. 1973. *Bayesian Inference in Statistical Analysis*. Wiley.
- [6] Good, J. 1965. *The Estimation of Probabilities: An essay on modern Bayesian methods*. MIT Press, Cambridge, US-MA.
- [7] Mankins, J. C. 1995. "Technology Readiness Levels." NASA.
- [8] Assistant Secretary of Defense for Research and Engineering (ASD(R&E)). 2011. Department of Defense Technology Readiness Assessment (TRA) Guidance, April.
- [9] Sauser, B., J. Ramirez-Marques, R. Magnaye, and W. Tan. 2008. "A Systems Approach to Expanding the Technology Readiness Level within Defense Acquisition." International Journal of Defense Acquisition Management Vol. 1: 39-58.
- [10] Austin, M. F., and D. M. York. 2015. "System Readiness Assessment (SRA) – An Illustrative Example." 13th Annual Conference on Systems Engineering Research (CSER), Hoboken, US-NJ, 18 – 20 March.

ABOUT THE AUTHORS

[Editor: Author biographies were current when the paper was initially published in 2017.]

Mr. Marc F. Austin is the technical director of a systems engineering group responsible for advancing systems engineering principles and practices for future implementation across the US Department of Defense. Mr. Austin holds a bachelor of science degree in electrical engineering, a master of science degree in applied mathematics and a certificate of special study in systems engineering from the Stevens Institute of Technology.

Dr. Cheyne Homberger holds a PhD in mathematics from the University of Florida, and conducts research in the areas of analytic and enumerative combinatorics. He currently serves as adjunct faculty at the University of Maryland, Baltimore County.

Ms. Virginia Ahalt holds an MS in statistics and operations research from the University of North Carolina, Chapel Hill. She currently works as an operations research analyst.

Ms. Erin Doolittle is a PhD candidate at Clemson University. She received an MS from Clemson University in mathematical sciences with a focus in operations research. Her research interests include multi-criteria optimization under conditions of uncertainty. She is currently employed as an operations researcher with the Department of Defense.

> continued on page 60

A Bayesian Approach for Estimating Complex System Reliability

Ozge Doguc and **Jose Emmanuel Ramirez-Marquez**

■ **ABSTRACT**
Although a number of recent studies on using Bayesian Networks (BN) for system reliability estimation have been proposed, these studies are based on the assumption that a pre-built BN was designed to represent the system. In these studies, the task of building the BN is typically left to a group of specialists who are BN and domain experts. However, the process of building a system-specific BN is generally very time consuming and may lead to incorrect deductions. As there are no existing studies to eliminate the need for a human expert in the process of system reliability estimation, this paper introduces a holistic method that uses historical data about the system to be modeled as a BN and provides efficient techniques for automated construction of the BN model and estimation of the system reliability. Moreover, very limited human intervention is sufficient for the process of BN construction and reliability estimation.

INTRODUCTION

System reliability can be defined as the probability that a system will perform its intended function during a specified period of time under stated conditions (Gran and Helminen 2001). Traditionally, engineers estimate reliability by understanding how the different components in a system interact for system success. Based on this understanding, typically a graphical model (usually in the form of a fault tree, a reliability block diagram, or a network graph) is constructed to represent component interactions. Using the graphical model, different analysis methods such as minimal cut sets, minimal path sets, Boolean truth tables, etc (Coyle, Arno, and Hale 2002; Fenton, Krause, and Neil 2002; Gopal, Kuolung, and Nader 2001) are used to represent system reliability quantitatively. At the end, the reliability characteristics of the components in the system are introduced into the mathematical representation in order to obtain a system level reliability estimate. This approach is valid whenever the system success or failure behavior is well understood. However, for complex systems (that is, systems with large numbers of components and/or complex component interactions), understanding component interactions, which usually requires intervention of a domain expert, may prove to be a challenging problem.

Bayesian networks (BN) have been proposed as an alternative to traditional reliability estimation approaches (Amasaki *et al.* 2003; Boudali and Dugan 2006; Gran and Helminen 2001). BN have significant advantages over traditional frameworks, partly because they are easy to use in interaction with domain experts in the reliability field (Sigurdsson, Walls, and Quigley 2001). Current approaches for reliability analysis via a BN (Amasaki *et al.* 2003; Bobbio *et al.* 2001; Sigurdsson, Walls, and Quigley 2001) use specialized networks, each of which is designed for a specific system. In these studies, the BN structure to be used for estimating system reliability should be known *a priori*. This assumption presupposes that the BN should be built by an expert who has "adequate" knowledge about the system behavior.

However, finding such an expert may not be possible at all times for every system under consideration. Moreover, the number of such experts is limited and finding one is usually difficult and costly (Lagnseth and Portinale 2005). Also, human intervention is always open to unintentional mistakes, which could cause discrepancy in the results. These issues are particularly true in complex systems, where the number of components and interactions are larger and thus, the likelihood of miscalculations can be substantial.

To address these issues, this study introduces a holistic method for estimating system reliability by linking BN construction from raw component and system data, association rule mining, and evaluation of conditional probabilities. Based on our literature review, this is the first study that incorporates these methods for estimating system reliability to reduce the need for human intervention. The proposed method automates the process of BN construction by using the *K2* algorithm (a commonly used association rule mining algorithm), which has been proven to be efficient and accurate for finding associations (Cooper and Herskovits 1992) from a dataset of historical data about the system. Moreover, unlike previous approaches, the proposed solution is not system specific, it can be applied to systems following any kind of configuration (two terminal, k-terminal, all terminal, etc.) and behavior (binary, capacitated, and multi-state). In essence, our approach can build a BN and estimate reliability for any system when observed system data is available (Doguc and Ramirez-Marquez 2007).

LITERATURE SURVEY

Estimating systems reliability using BN dates back as early as 1988, when it was first defined by Barlow (1988). The concept of BN has been discussed in several earlier studies (Cowell *et al.* 1999; Jensen 2001; and Pearl 1988); the idea of using BN in systems reliability has gained acceptance within the last decade because of the simplicity it allows to represent systems and the efficiency for evaluating component associations. More recently, BN have found applications in software reliability (Fenton, Krause, and Neil 2002; Gran *et al.* 2000), fault finding systems (Jensen 2001), and general reliability modeling (Bobbio *et al*. 2001). In recent studies, predefined BN are used for reliability estimation for specific systems. For example, Gran and Helminen (2001) study on building BN for nuclear power plants and introduce a hybrid method for estimating the reliability of the plant. In their study, they considered the nuclear plant as two subsystems: a software system and the plant hardware. Therefore, they combined two BN that were being used for corresponding systems: 1) The Halden Project (HRP) (Dahll and Gran 2000) uses a BN for risk assessment based on disparate evidences. 2) The VTT Automation (Helminen 2000) focuses on the reliability of software-based systems using BN. Additionally they discuss another challenge; each BN uses a different modeling and simulation environment.

In another study Helminen and Pulkkinen present a BN-based method for reliability estimation of computer-based motor protection relay (Helminen and Pulkkinen 2003). In their study, Helminen and Pulkkinen assume existence of a BN that models the system and introduce methods for estimating prior probabilities and assessing the system reliability accordingly.

In addition to these, Amasaki *et al.* (2003) use BN for software quality assessment. They modeled the phases of a software system as a BN, and by using this model they simulated the faults that may occur in their system. After this step, they used the actual data and performed sensitivity analysis of the BN model that they constructed. In addition to these, Boudali and Dugan (2006) introduce a method for reliability assessment in dynamic systems by using temporal BN; where the system components change states at different time intervals. Moreover, Singh *et al.* (2001) presents their work on reliability estimation in component based systems. They classify the component-based system reliability estimation methods into three as state based models, path based models, and additive models.

Although all of the studies introduced in this section use BN for reliability estimation, they require human domain experts to evaluate the prior probabilities and understand the structure of the BN. In the next section, we introduce a methodology that automates the process of BN construction and reduces the need for a human expert for system reliability estimation.

BAYESIAN NETWORKS

As discussed in the previous sections, BN have been used in various studies for estimating system reliability. In this section we first provide definitions of BN and Bayes theorem. Then we discuss the *K2* algorithm that we used to create BN in this study.

Using Bayesian Networks for System Reliability. One could summarize the BN as an approach that represents the interactions between the variables from a probabilistic perspective. This representation is modeled as a directed acyclic graph, where the nodes represent the variables and the links between each pair of nodes represent the causal relationships between the variables. In general, a fundamental assumption for the construction of a BN is that the strength of the interaction/influence among the graph nodes is uncertain and thus, this uncertainty is represented by assigning a probability of existence to each of the links between nodes.

From systems engineering perspective, the variables of a BN are defined as the components in the system while the links represent the interactions of the components leading to system "success" or "failure." Under a reliability analysis perspective, a variable A in BN constitutes the success of a specific system component and therefore, p(A) represents the probability of success for such a component. For non-trivial systems – systems not following a series, parallel or any combination of these configurations – the failure/success probability of a system is usually dependent on the failure/success of a non-evident collection of components. Strictly speaking, the probability of success of a component is conditional on the available evidence from other components. In a BN this dependency is represented as a directed link between two components, forming a *child* and *parent* relationship, so that the dependent component is called as the *child* of the other. Therefore, the success probability of a child node is *conditional* on the success

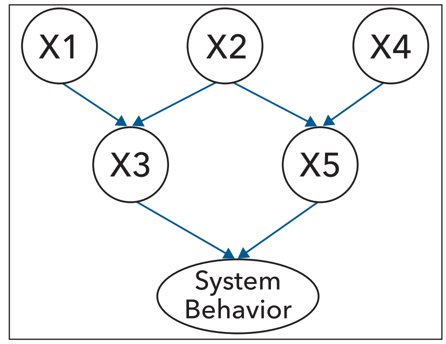


Figure 1. A sample Bayesian network

probabilities associated with each of its parents (Fenton, Krause, and Neil 2002). The *conditional probabilities* of the child nodes are calculated by using the Bayes' theorem via the probability values assigned to the parent nodes. Also, absence of a link between any two nodes of a BN indicates that these components do not directly interact for system failure/success thus, they are considered *independent* of each other and their probabilities are calculated separately.

To illustrate these concepts, the BN shown in Figure 1 presents how five components of a system interact. In this BN the child-parent relationships of the components can be observed, where on the quantitative side the *degrees* of these relationships (associations) are expressed as probabilities (Lagnseth and Portinale 2005).

In Figure 1 the topmost nodes (X_1, X_2, and X_4, representing components 1, 2, and 4 respectively) do not have any incoming edges, therefore they are conditionally independent of the rest of the components in the system. The *prior probabilities* that are assigned to these nodes should be known beforehand—with the help of a domain expert or using historical data about the system. Based on these prior probabilities, the conditional probability table (CPT) that belong to a *dependent* node, such as X_3, can be calculated using Bayes' theorem as illustrated by equation (1):

$$p(X_3 \mid X_1, X_2) = \frac{p(X_1, X_2 \mid X_3)\, p(X_3)}{p(X_1, X_2)} \quad (1)$$

Equation (1) shows that the probability for the node X_3 is *independent* of nodes other than X_1 and X_2 in the system. Similar to prior probabilities, CPT can be computed by using historical system and component data. However, an important question on how to discover the associations among the system components still remains. As an alternative to using a domain expert for this purpose, an unsupervised BN construction algorithm, *K2* is used in this paper.

The *K2* Algorithm. The *K2* algorithm, for construction of a BN, was first defined

Table 1. Dataset for the illustrative example

Observation	X_1	X_2	X_3	X_4	X_5	*System Behavior*
1	1	1	0	0	0	1
2	0	1	1	0	0	0
3	1	0	1	1	1	1
4	0	0	0	0	0	0
5	1	1	1	0	1	1
6	0	1	1	1	0	0
7	1	0	0	1	0	1
8	0	0	1	1	1	1
9	1	1	1	0	0	0
10	0	1	0	1	1	1

by Cooper and Herskovits (1992) as a greedy heuristic search method. This algorithm searches for the parent set for a node that has the maximum association with it. The *K2* algorithm is composed of two main factors: a scoring function *f* to quantify the associations and rank the parent sets according to their scores, and a heuristic to reduce the search space to find the parent set with highest degree of association. The *K2* algorithm would need to examine all possible parent sets, i.e., starting from the empty set, it should consider all subsets of set of possible parents without the heuristic. With the help of the heuristic, the *K2* algorithm does not need to consider the whole search space; it starts with the assumption that the node has no parents and adds incrementally that parent whose addition most increases the scoring function. The *K2* algorithm stops adding parents to the node when addition of no single parent can increase the score.

ILLUSTRATION OF OUR METHODOLOGY

This section provides a step-by-step explanation of the BN construction framework and system reliability estimation method discussed in the previous section. Table 1 presents an example historical dataset that contains observations on the sample system shown in Figure 1 with five components labeled X_1 to X_5. Each row in Table 1 shows the states of the system components at an instance of time t_i; when the observation was done. For the sake of simplicity and without loss of generality in the proposed method, component failure data exhibits binary behavior.

That is, for each component X_i, the value of 0 represents failure while the value of 1 represents full functionality for the corresponding observation. Also, in Table 1, information about the overall *System Behavior* is provided in last column.

Our proposed method uses a dataset such as displayed in Table 1; finds associations between the columns (system components); calculates the degrees of these associations; builds the associated BN and finally uses it to estimate overall system reliability. In the first step of our method the *K2* algorithm starts with the first component in the dataset, X_1. Since X_1 does not have any succeeding components (i.e., possible candidate parents), the *K2* algorithm skips it and picks the second component in the dataset, which is X_2.

For *X*2, there are two alternative parent sets: the empty set φ, or X_1. Therefore, the *K2* algorithm computes the scoring function f for each of these alternative parent sets and compares the results. Then, the set of candidate parents with highest *f* score is chosen as the parent set for X_2. At the end of this iteration the values $\frac{1}{2310}$ and $\frac{1}{3600}$ are calculated and then compared; and the former, representing the score of the empty set {φ}, picked as the parent. So the *K2* algorithm decides that X_2 has no parents, which means that there is no association between X_1 and X_2.

In the next iterations of the *K2* algorithm, the number of possible candidate parent sets to be considered and the number of computations for *f* score calculation increases. Skipping the details, *f* scores of the candidate parent sets for the X_3 component are given in Table 2. Because the *K2* algorithm iterates on the components according to their ordering in dataset, components X_4 and X_5 are not taken into account as candidate parents for X_3. The *K2* algorithm selects the set $\{X_1, X_2\}$ as parent set of X_3, because it has the highest *f* score. The number of computations grows with the order of the component in the system, and when the *K2* algorithm finishes with the last column (*System Behavior* in Table 1), it outputs the BN structure displayed in Figure 1.

The next step of the proposed method estimates system reliability using the BN that was constructed by the *K2* algorithm.

Table 2. f scores for all possible candidate parent sets for X_3

Parent Set	*f* score
φ	$\frac{1}{2310}$
$\{X_1\}$	$\frac{1}{2772}$
$\{X_2\}$	$\frac{1}{3600}$
$\{X_1, X_2\}$	$\frac{1}{288}$

Besides the associations that were discovered by the *K2* algorithm in the previous step, the inference rules should be used to calculate the conditional probabilities between the nodes in the BN. The conditional probabilities are essential in calculating the overall reliability of the system, as they represent the degrees of associations between components of a system. Each component with a non-empty parent set in the network is associated with a CPT. In this step, with the help of CPT and the prior probabilities that X_1 and X_2 have, the success probability value for X_3 can be calculated. According to the BN structure in Figure 1, components *X*1 and *X*2 are independent of others; therefore their success probabilities can be directly inferred from the observations dataset in Table 1. From Table 1 it can be evaluated that $p(X_1 = 1) = 0.5$ and $p(X_2 = 1) = 0.6$ and the probability of success for component X_3 can be calculated as 0.57 using Bayes' rule provided in Equation (1). Extending the computations for the other components in the network, success probabilities for the rest of the components in the sample system can be evaluated; such that $p(X_4 = 1) = 0.4$ and $p(X_5 = 1) = 0.6$. In the last step, the system reliability can be calculated by using these probability values and the CPT of the "system behavior" node in the BN structure given in Figure 1. The success probability for the *system behavior* node is calculated as 0.72 or 72%; which is the reliability of the sample system presented in this section. The proposed method for estimating system reliability using observations dataset is superior to previously defined methods due to its unsupervised nature; almost all steps of the required computations can be carried out without any human intervention.

EXPERIMENTAL ANALYSIS

In this section, experimental analysis on the performance of our proposed method for system reliability estimation is provided. In order to give a better perception of analysis, performances of the two phases of the proposed method (BN construction and reliability estimation) are examined separately. First, performance and correctness of the *K2* BN construction algorithm is analyzed using historical data (obtained via Monte Carlo simulation) for the following BN.

BN displayed in Figure 2 represent different systems with various components. For our experimental analysis, separate data sets – similar to Table 1 – are used for each example BN. As it was explained in the previous section, the *K2* algorithm uses historical system data as input.

Therefore running time of this algorithm is highly dependent on the size of the input

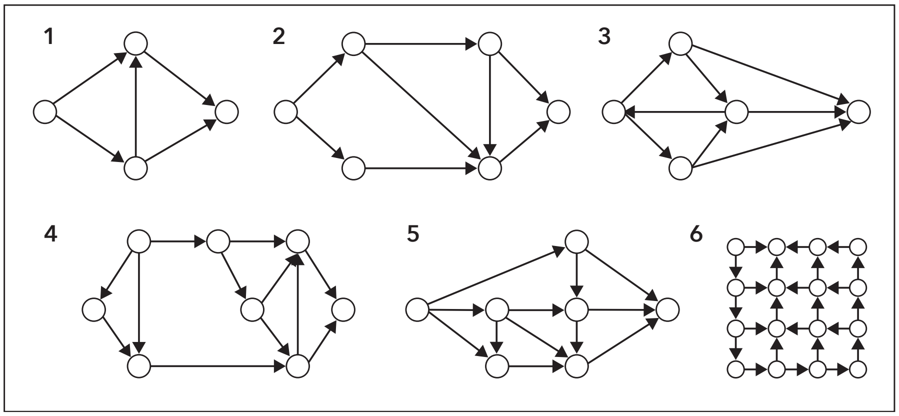


Figure 2. Case BN tested on the K2 algorithm

data set, that is, number of nodes (n) and number of observations (t). Each of the case BN shown in Figure 2 has different number of nodes ranging from 5 to 16, and the performance of the *K2* algorithm on each BN is analyzed using different input data sets. Figure 3 shows the experimental results on the performance of the *K2* algorithm. It can be observed from Figure 3 that the running time of the *K2* algorithm is quadratic ($O(n^2)$) with the number of nodes and linear with the number of observations. This is an expected result, since the *K2* algorithm reduces the time-complexity of finding associations from exponential (2^n) to quadratic (n^2). This brings the conclusion that for even substantially large systems ($n > 100$) the *K2* algorithm will be efficient to use (that is, doing 10,000 iterations instead of 2^{100}).

The number of observations that are used for discovering the associations between system components is an important measure for both efficiency of the *K2* algorithm and correctness of the constructed BN. Also accuracy of the reliability estimation is highly dependent on correctness of the underlying BN model. Errors in the *K2* algorithm would lead to incorrect assignments of associations in the BN; which will end up with inaccuracies in the reliability values. Once the BN is correctly constructed, estimating the system reliability is simple and straightforward as discussed in the previous section; therefore, in this section we evaluate correctness of the BN constructed by the *K2* algorithm. The *K2* algorithm can be expected to find out associations more accurately when more observations used as input (Cooper and Herskovits 1992). Using the constructed BN, error rate (ρ) of the *K2* algorithm can be calculated by using Equation (2):

$$\rho = \frac{A_M}{A_T} = \frac{A_{FP} + A_{FN}}{A_T} \qquad (2)$$

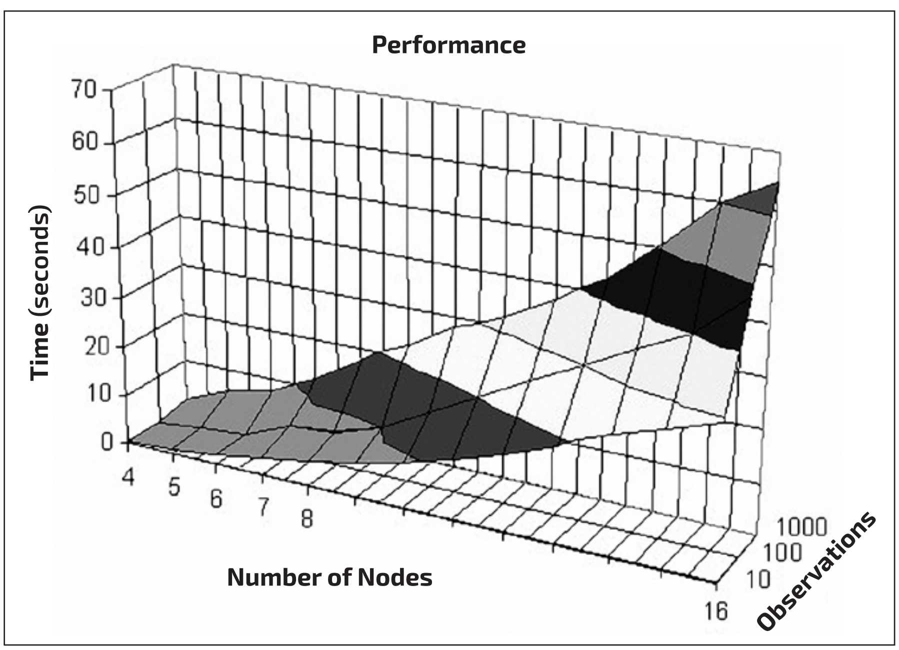


Figure 3. Running time of the K2 algorithm

In Equation (2), a *false positive* (A_{FP}) is defined as an association decided by the *K2* algorithm; however does not exist in the actual BN given in Figure 2. Conversely, a *false negative* (A_{FN}) is defined as an existing association in the actual BN that is missed by the *K2* algorithm. Both should be taken into account while calculating accuracy (where accuracy = $1-\rho$) of the *K2* algorithm with the constructed BN. In this study, correctness of the constructed BN models are evaluated by using data sets with 10, 100 and 1000 observations for different case networks. General analysis results on the correctness and accuracy are provided in Figure 4.

According to Figure 4, regardless of the size of the constructed BN, accuracy of the BN model increases as more observations are used. This is an expected result, since associations between the components are decided by using the observations and these associations can be figured out more precisely when more observations available.

We used the case networks displayed in Figure 2 in our experiments. Our experimental results for the accuracy of BN construction as well as the CPU times for BN construction and reliability estimation are presented in Table 3. For our experiments we used a computer equipped with an Intel Centrino 2Ghz CPU and 2GB RAM. Moreover, we implemented our proposed method in Matlab 7.0.

CONCLUSIONS

Estimating system reliability using BN is a very popular practice and has been widely studied recently. There are numerous methods in the literature defined for estimating system reliability, which are mainly focused on doing it for specific systems, such as nuclear plants. However, none of these studies dealt with the problem of requiring a human expert to construct the BN. This is the first study that introduces a methodology for efficient construction of BN models and estimating system reliability, with limited very human expert requirement. The proposed method uses historical data about the system to be modeled and constructs the BN model automatically. The *K2* algorithm is used for this purpose, which is a popular and efficient association rule mining method.

Next it was shown that the system reliability can efficiently be estimated by using the BN model. According to the experimental results, reducing the running time of finding associations from O(*2n*) to O(*n2*), the proposed methodology can work efficiently even with substantially large systems. Moreover, the BN models constructed by the *K2* algorithm are shown to be accurate, especially when more

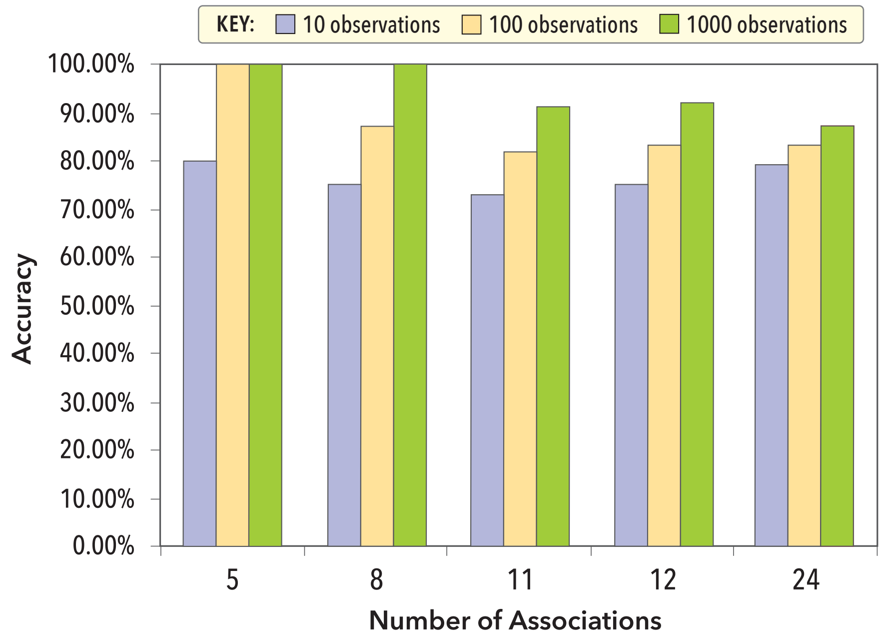


Figure 4. Accuracy of results in BN construction

Table 3. Compilation of results for case BN

Case Network	Number of Nodes	Number of associations	*K2* algorithm CPU time (seconds)	Accuracy of the constructed BN (w/1000 observations)	Reliability Estimation CPU time (seconds)
1	4	5	1.839	100.00%	0.465
2	6	8	10.471	100.00%	0.981
3	5	8	5.567	100.00%	0.830
4	8	11	37.898	90.91%	1.004
5	7	12	26.012	91.66%	0.991
6	16	24	148.762	87.50	1.587

historical data about the system is available. As expected, the experimental results show that when 1,000 historical observations on the system are available, the constructed BN are more than 90% accurate. Accuracy of the constructed BN is highly influential on the correctness of the system reliability values, as incorrect associations in the BN would lead to biased calculations while estimating system reliability. In conclusion, the methodology introduced in this study will help systems engineers as it minimizes human interaction and provides efficient ways of automatically building a BN model and estimating system reliability. ■

REFERENCES

- Amasaki, S., Y. Takagi, O. Mizuno, and T. Kikuno. 2003. "A Bayesian Belief Network for Assessing the Likelihood of Fault Content." In *Proceedings of the 14th International Symposium on Software Reliability Engineering.*
- Barlow, R. E. 1988. "Using influence diagrams." *Accelerated life testing and experts' opinions in Reliability*: 145–150.
- Bobbio, A., L. Portinale, M. Minichino, and E. Ciancamerla. 2001. "Improving the analysis of dependable systems by mapping fault trees into Bayesian networks." *Reliability Engineering and System Safety* 71 (3): 249–260.
- Boudali, H., and J. B. Dugan. 2006. "A Continuous-Time Bayesian Network Reliability Modeling, and Analysis Framework." *IEEE Transaction on Reliability* 55 (1): 86-97.
- Cooper, G. F., and E. Herskovits. 1992. "A Bayesian method for the induction of probabilistic networks from data." *Machine Learning* 9 (4): 309-347.
- Cowell, R. G., A. P. Dawid, S. L. Lauritzen, and D. J. Spiegelhalter. 1999. *Probabilistic networks and expert systems*. New York, US-NY: Springer-Verlag.
- Coyle, T., R. G. Arno, and P. S. Hale. 2002. "Application of the minimal cut set reliability analysis methodology to the gold book standard network." In *Proceedings of the Industrial and Commercial Power Systems Technical Conference*.
- Dahll, G., and B. A. Gran. 2000. "The Use of Bayesian Belief Nets in Safety Assessment of Software Based Systems." *Special Issues of International Journal on Intelligent Information Systems* 24 (2): 205-229.
- Doguc, O., and J. E. Ramirez-Marquez. 2007. "A Generic Method for Estimating System Reliability Using Bayesian Networks." *Journal of Reliability Engineering and System Safety (under revision)*.
- Fenton, N., P. Krause, and M. Neil. 2002. "Software Measurement: Uncertainty and Causal Modeling." *IEEE Software* 10 (4): 116-122.
- Gopal, C., H. Kuolung, and A. Nader. 2001. "A New Approach to System Reliability." *IEEE Transactions on Reliability* 50 (1): 75-84.
- Gran, B. A., G. Dahll, S. Eisinger, E. J. Lund, J. G. Norstrøm, P. Strocka, and B. J. Ystanes. 2000. "Estimating Dependability of Programmable Systems Using BBNs." In *Proceedings of the Safecomp 2000*.
- Gran, B. A., and A. Helminen. 2001. "A Bayesian Belief Network for Reliability Assessment." *Safecomp 2001* 2187: 35–45.
- Helminen, A. 2000. Reliability Estimation of Software-based Digital Systems Using Bayesian Networks. Espoo: Helsinki University of Technology.
- Helminen, A., and U. Pulkkinen. 2003. "Quantitative Reliability Estimation of a Computer-Based Motor Protection Relay Using Bayesian Networks." *Safecomp 2003* 2788: 92–102.
- Jensen, F. V. 2001. *Bayesian networks and decision graphs*. New York, US-NY: Springer Verlag.
- Lagnseth, H., and L. Portinale. 2005. "Bayesian Networks in Reliability."
- Pearl, J. 1988. *Probabilistic Reasoning in Intelligent Systems*. San Francisco, US-CA: Morgan Kaufmann.
- Sigurdsson, J. H., L. A. Walls, and J. L. Quigley. 2001. "Bayesian belief nets for managing expert judgment and modeling reliability." *Quality and Reliability Engineering International* 17: 181–190.
- Singh, H., V. Cortellessa, B. Cukic, E. Gunel, and B. Bharadwaj. 2001. "A Bayesian Approach to Reliability Prediction and Assessment of Component Based Systems." In *Proceedings of the 12th International Symposium on Software Reliability Engineering* (*ISSRE'01*).

ABOUT THE AUTHORS

[Editor: Author biographies were current when the paper was initially published in 2009.]

Ozge Doguc is a business analyst at Asset Control Inc. in New York, NY. Prior to that, she had worked as a research assistant in the Systems Engineering Department at Stevens Institute of Technology between 2004 and 2006. She holds a ME degree in systems engineering from Stevens Institute of Technology. She graduated from Beykent University in Istanbul, Turkey obtaining BS degrees in mathematics and computer sciences, and management information systems. She is currently a PhD candidate in systems engineering and her research interests and areas include reliability estimation in complex systems, grid systems and system operational effectiveness.

Jose E. Ramirez-Marquez is an assistant professor at the Stevens Institute of Technology in the Department of Systems Engineering. His research interests include system reliability and quality assurance, uncertainty modeling, advanced heuristics for system reliability analysis, applied probability and statistical models and applied operations research. He has authored more than 12 articles in leading technical journals on these topics. He obtained his PhD at the Rutgers University in industrial and systems engineering. He received his BS degree in actuarial science from the UNAM in Mexico City in 1998 and MS degrees in industrial engineering and statistics from the Rutgers University.

Salado and **Nilchiani** continued from page 46

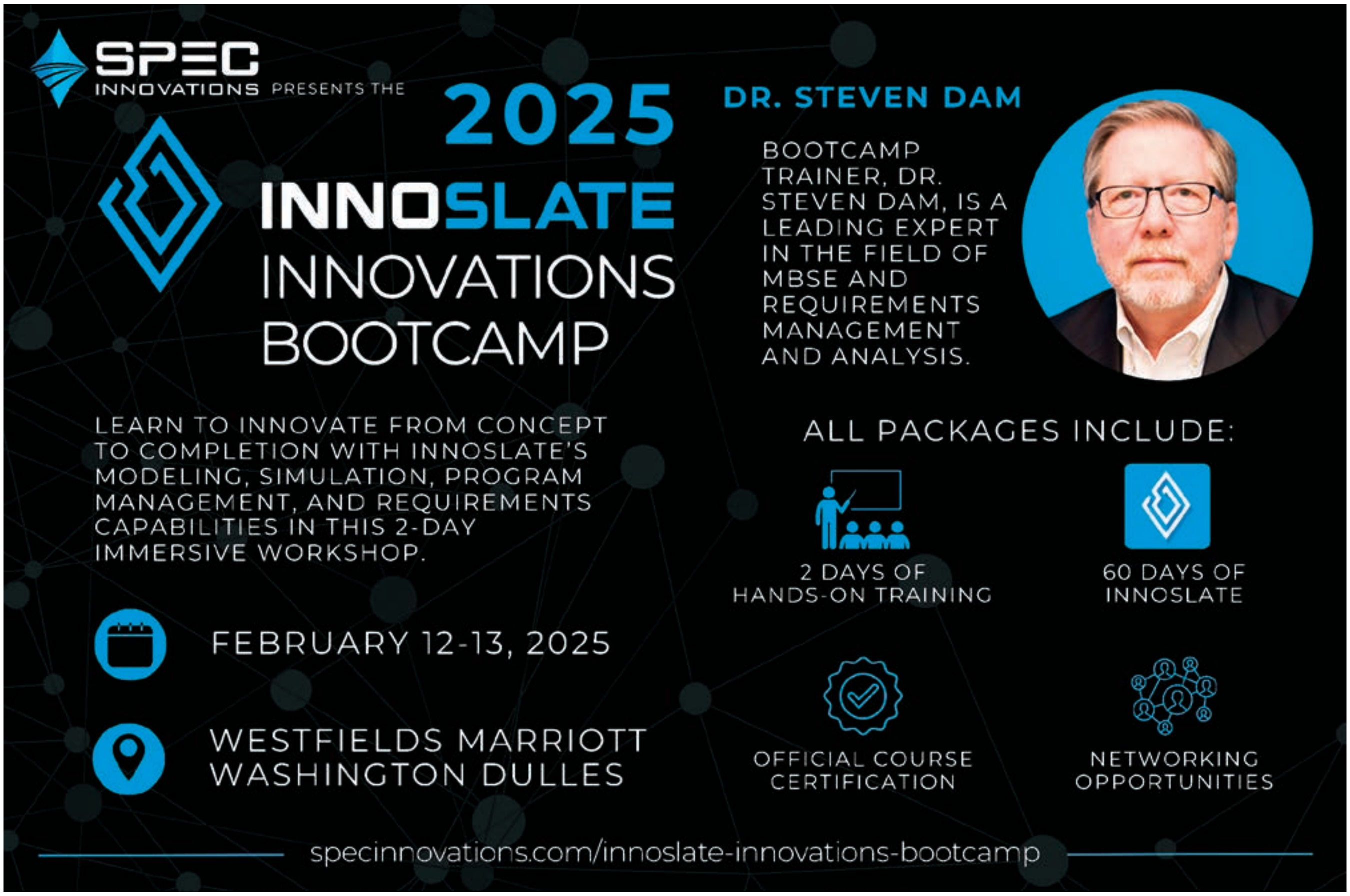
Dr. Roshanak Nilchiani is an assistant professor of systems engineering in the School of Systems and Enterprises at Stevens Institute of Technology. She has joined Stevens Institute of Technology as an assistant professor in 2006 and since then, she has been focusing on two major tracks of research, the first on focusing on flexible and adaptive designs and measuring the value of various systems illities. The second track of her current research focuses on quantifying, measuring, and embedding Resilience and sustainability in large scale critical infrastructure systems. Dr. Nilchiani received her PhD in aerospace systems from Massachusetts Institute of Technology in 2005, where she focused on designing, embedding, and measuring of flexibility in space systems and spacecrafts. She has used various decision-making tools, uncertainty modeling, options and decision analysis tools and integrated them into her models for measuring the value of flexibility and she has been extensively using various economics tools for capturing cost and various types of benefits in monetary and non-monetary utility forms. Her research has been funded by DARPA's F6 program and Department of Homeland Security (DHS) center of Excellence on Resilient Infrastructure systems.

Austin et al. continued from page 54

Dr. George Polacek is a senior technical advisor for enterprise systems engineering and architecture in the Department of Defense. He holds a PhD in systems engineering from Stevens Institute of Technology. His research interest is the application of self-organizing mechanisms found in biological and social systems to the engineering of autonomous systems, systems-of-systems, and enterprises. Dr. Polacek also holds a BS in electrical engineering from the University of Illinois, an MS in electrical engineering from Johns Hopkins University, and is a member of INCOSE.

Dr. Donald M. York is a technical fellow and senior systems engineer at Engility Corporation. Dr. York has 38 years of systems engineering and technical leadership experience in advanced technology systems. He holds a PhD in systems engineering and both bachelor's and master's degrees in mechanical engineering. Dr. York is an INCOSE certified systems engineering professional and a licensed professional engineer.


2025
Annual INCOSE
international workshop
HYBRID EVENT
Seville, SPAIN
February 1 - 4, 2025
MARK YOUR CALENDAR!
The INCOSE International
Workshop is traveling to
SEVILLE
February 1 - 4, 2025
https://www.incose.org/iw2025


SPEC INNOVATIONS PRESENTS THE
2025
INNOSLATE
INNOVATIONS
BOOTCAMP
DR. STEVEN DAM
BOOTCAMP TRAINER, DR. STEVEN DAM, IS A LEADING EXPERT IN THE FIELD OF MBSE AND REQUIREMENTS MANAGEMENT AND ANALYSIS.
LEARN TO INNOVATE FROM CONCEPT TO COMPLETION WITH INNOSLATE'S MODELING, SIMULATION, PROGRAM MANAGEMENT, AND REQUIREMENTS CAPABILITIES IN THIS 2-DAY IMMERSIVE WORKSHOP.
ALL PACKAGES INCLUDE:
2 DAYS OF HANDS-ON TRAINING
60 DAYS OF INNOSLATE
FEBRUARY 12-13, 2025
WESTFIELDS MARRIOTT WASHINGTON DULLES
OFFICIAL COURSE CERTIFICATION
NETWORKING OPPORTUNITIES
specinnovations.com/innoslate-innovations-bootcamp


INCOSE VOLUNTEER OPPORTUNITY
THE BEST ENGINEERS ALLOW FOR A LITTLE GIVE.
Become an INCOSE volunteer today!
incose.org/volunteer
INCOSE
A better world through a systems approach


CATIA MAGIC
DESIGN AND ENGINEER FASTER WITH THE CLOUD-BASED COLLABORATIVE MBSE SOLUTION
LEARN MORE
DASSAULT SYSTEMES