



# Interactive Model-Centric Systems Engineering (IMCSE)

## Phase 3

Technical Report SERC-2015-TR-043-1

March 1, 2016

### Investigators

**Principal Investigator:** Dr. Donna H. Rhodes, Massachusetts Institute of Technology

**Co-Investigator:** Dr. Adam M. Ross, Massachusetts Institute of Technology

### Phase 3 Research Team

Dr. Adam Ross, Massachusetts Institute of Technology

Dr. Donna Rhodes, Massachusetts Institute of Technology

Copyright © 2016 Stevens Institute of Technology

The Systems Engineering Research Center (SERC) is a federally funded University Affiliated Research Center managed by Stevens Institute of Technology.

This material is based upon work supported, in whole or in part, by the U.S. Department of Defense through the Office of the Assistant Secretary of Defense for Research and Engineering (ASD(R&E)) under Contract HQ0034-13-D-0004 (TO 0003).

Any views, opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the United States Department of Defense nor ASD(R&E).

**No Warranty.**

This Stevens Institute of Technology and Systems Engineering Research Center Material is furnished on an “as-is” basis. Stevens Institute of Technology makes no warranties of any kind, either expressed or implied, as to any matter including, but not limited to, warranty of fitness for purpose or merchantability, exclusivity, or results obtained from use of the material. Stevens Institute of Technology does not make any warranty of any kind with respect to freedom from patent, trademark, or copyright infringement.

This material has been approved for public release and unlimited distribution.

## Contents

<b>Figures .....</b>	<b>5</b>
<b>Executive Summary of Phase 3 .....</b>	<b>7</b>
<b>Purpose of Research .....</b>	<b>7</b>
<b>Work Accomplished in Phase 3.....</b>	<b>7</b>
<b>Research Results.....</b>	<b>9</b>
<b>Introduction .....</b>	<b>10</b>
<b>Motivation.....</b>	<b>10</b>
Insufficiencies in Current Practice.....	10
<b>Project Overview: IMCSE .....</b>	<b>11</b>
<b>What is IMCSE? .....</b>	<b>11</b>
<b>Research Program Vision.....</b>	<b>12</b>
<b>IMCSE Pillars – Four Topic Areas .....</b>	<b>12</b>
Big Data .....	12
Visual Analytics.....	13
Complex Systems .....	14
Model-Based Systems Engineering (MBSE) .....	17
Challenges across the Pillars .....	17
<b>Research Progress .....</b>	<b>19</b>
<b>IMCSE Research Roadmap.....</b>	<b>19</b>
Developing the Research Roadmap .....	19
<b>Investigation of Selected Research Themes.....</b>	<b>21</b>
Digital System Model .....	21
Model Curation .....	34
<b>Cognitive and Perceptual Considerations in Human-Model Interaction.....</b>	<b>42</b>
Background.....	42
Glass Cockpit Investigation .....	42
Cognitive & Perceptual challenges.....	47
Discussion.....	49
Next Steps .....	50
<b>Interactive Epoch-Era Analysis.....</b>	<b>51</b>
Background.....	51
A Framework for IEEA .....	53
Demonstration Case: Space Tug.....	57
Discussion.....	74
Next Steps .....	75
<b>Model Choice and Tradeoff .....</b>	<b>76</b>
Background.....	76
Demonstration of Evaluative Model Trading: Space Tug .....	78
Preliminary Demonstration of Combined Value Model and Evaluative Model Trading: Space Tug .	88
Next Steps .....	109
<b>Synthesis of IMCSE Research Progress .....</b>	<b>110</b>
<b>Moving Forward to Phase Four .....</b>	<b>113</b>

<b>References .....</b>	<b>114</b>
-------------------------	------------

## FIGURES

---

Figure 1. Five Aspects of Complex Systems <sup>12</sup> .....	16
Figure 2. Research at the Intersection of four pillars .....	18
Figure 3. Four Focus Areas of IMCSE Research Roadmap .....	20
Figure 4. Importance of a Model Curation Function .....	37
Figure 5. Shift from the ‘workbench’ to model-centric environment .....	42
Figure 6. Displays used by pilot (a) traditional, analog cockpit; (b) glass cockpit. ....	44
Figure 7. Interactive Epoch-Era Analysis processes and modules .....	54
Figure 8. Interactive Epoch-Era Analysis modules.....	55
Figure 9. Single attribute utility (SAU) curves for the 8 different missions.....	59
Figure 10. Prototype Application for Interactive Visualization of MAU weightings showing (a) Initial tradespace and preferences over performance attributes; (b) Revised preferences and final tradespace.....	59
Figure 11. Combinations of Design-Epoch-Era evaluated through performance, value, context and strategy models .....	61
Figure 12. Interactive Filtering Application .....	63
Figure 13. Interactive heatmap visualization.....	63
Figure 14. Interactive Filtering Application implementing OLAP. ....	64
Figure 15. (a) Interactive force-directed graphs of change mechanisms between designs; (b) Betweenness centrality metric assigned to node radius.....	69
Figure 16. Interactive application showing candidate designs across a single era (a) before filtering; and (b) after.....	70
Figure 17. Visualization showing eras visualized using parallel sets .....	72
Figure 18. Interactive chord diagram used to visualize design change behavior.....	73
Figure 19. Role of key models for supporting system decision making, with alternative value models use in demonstration case .....	77
Figure 20. Role of key models in system decision making, with example evaluative models called out.....	79
Figure 21. Space Tug model implementation #1 and resulting tradespace.....	81
Figure 22. Space Tug model implementation #2 and resulting tradespace.....	82
Figure 23. Space Tug model implementation #3 and resulting tradespace.....	82
Figure 24. Space Tug model implementation #4 and resulting tradespace.....	83

Figure 25. The tradespaces for all four implementations, with all four Pareto sets marked in each .....	84
Figure 26. The tradespaces for all four implementations, highlighting the model artifacts impacting designs 95/479 (magenta), combining to make them inefficient in implementation 4 only. Design numbers left/above of the points are on the Pareto front, right/below are inefficient. ....	86
Figure 27. The tradespaces for implementations 1 (left) and 4 (right), highlighting design variable patterns on the Pareto front. The six categories of designs in Table 12 are used as the color scheme.....	88
Figure 28. Four objective sets of two objectives each .....	90
Figure 29. Gridmap showing relative sizes of 0% fuzzy joint Pareto sets in evaluative model #191	
Figure 30. Gridmap showing relative sizes of 7% fuzzy joint Pareto sets in evaluative model #192	
Figure 31. Gridmap showing relative sizes of 0% fuzzy joint Pareto sets in evaluative model #293	
Figure 32. Gridmap showing relative sizes of 7% fuzzy joint Pareto sets in evaluative model #294	
Figure 33. Gridmap showing relative sizes of 0% fuzzy joint Pareto sets in evaluative model #396	
Figure 34. Gridmap showing relative sizes of 6% fuzzy joint Pareto sets in evaluative model #396	
Figure 35. Gridmap showing relative sizes of 0% fuzzy joint Pareto sets in evaluative model #498	
Figure 36. Gridmap showing relative sizes of 6% fuzzy joint Pareto sets in evaluative model #498	
Figure 37. Comparison of the payload capability single attribute utility function impact across the four evaluative models with MAU (blue triangle) and MOE (red triangle) Pareto sets indicated .....	100
Figure 38. Comparing evaluative model #1 first joint Pareto design at fuzzy level 7%(left) and evaluative model #4 first joint Pareto at fuzzy level 6% (right) illustrates how increase in maximum cost design impacts fuzzy metric.....	101
Figure 39. AHP tradespace scatterplot under the old speed models (#1 and #3, left) and new speed models (#2 and #4, right), with families of designs indicated by propulsion type (red=biprop/cryo, green=nuclear, blue=electric) .....	105
Figure 40. CBA tradespace scatterplot under the old speed models (#1 and #3, left) and new speed models (#2 and #4, right), with families of designs indicated by propulsion type (red=biprop/cryo, green=nuclear, blue=electric) .....	106
Figure 41. Comparison of MOE value model across evaluative models #1 and #2 (left) and models #3 and #4 (right).....	107
Figure 42. CBA value model scatterplot for evaluative model #4 (combined new material and new speed).....	108

## **EXECUTIVE SUMMARY OF PHASE 3**

---

This report is the technical report for Phase 3 of the Interactive Model-Centric Systems Engineering (IMCSE) research project. Portions of the Phase 1 and 2 reports have been repeated in this report for completeness.

---

### **PURPOSE OF RESEARCH**

IMCSE advances the current state of SE knowledge in “non-technical” aspects of model-based engineering. While MBSE and MBE activities are advancing technical aspects of models in the engineering of systems, this topic advances knowledge relevant to human interaction with models and model-generated information. It brings relevant knowledge from other fields (e.g., cognitive science, visual analytics, data science), placing it in context of systems engineering. Additionally, this research generates knowledge impacting human effectiveness in model-centric environments of the future including foundational theory, role of humans in designing/sustaining these environments and mitigating challenges rooted in cognitive and perceptual considerations.

IMCSE research advances knowledge concerning MPTs that enable reasoning, comprehension and collaborative decision making in the face of uncertainty, combining artificial and real data, and effectively utilizing vast amounts of information. New knowledge is created through normative and descriptive research approaches, leading to prescriptive outcomes that can be transitioned into practice, as well as educating the workforce. IMCSE research provides support to SERC’s four thematic research areas, and is most closely aligned with the Systems Engineering and Management Transformation in regard to decision-making capabilities and leveraging the capabilities of computation, visualization, and communication for quick and agile response. Decision-making will be increasingly informed by models, and this research topic strongly supports MPTs to enable this. There are many past and ongoing research topics that have relevance for the future of interactive model-centric environments and provide valuable input to this research.

---

### **WORK ACCOMPLISHED IN PHASE 3**

**1. Research Roadmap.** Researchers developed a research roadmap based on results from the pathfinder project to elicit information on the state of the IMCSE art and practice, with insights from the phase 2 workshop and additional meetings with research stakeholders. The research team clarified and expanded the urgent research questions, and investigated research priorities. In support of the roadmap development, two topics were explored in further detail.

The technical challenges for models are relatively well understood and there is significant progress being made in addressing them. The research team explored the non-technical challenges for the Digital System Model in context of the larger topic of the digital framework including Digital Thread and Digital Twin. Contributing to the continued understanding of problems and concerns, the research explored non-technical aspects, specifically intellectual property and knowledge assessment. As engineering practice becomes increasingly model-centric, models are valuable assets for designing and evolving systems. The second topic explored

was the need for model curation, which accordingly becomes a necessary functional role in organizations. The research team performed exploratory investigation to identify the various facets of a curation function, and the aspects for developing a shared understanding of models.

**2. Interactive Epoch-Era Analysis.** The researchers improved the approach for evaluating systems under dynamic uncertainty using epoch-era analysis, with focus on enhanced interactive capability and scaling for big data analysis. Feedback from practitioners and researchers validated the importance of continuing this research activity, which combines innovative methods with new enabling visualization and data handling technologies.

**3. Model Tradeoff and Choice.** A framework for conducting value model trades and evaluative (performance, cost) model trades was developed and tested to validate the framework and identify workflow considerations. The MIT Interactive Value-Driven Tradespace Exploration and Analysis Suite (IVTea Suite) was modified to support case demonstrations by applying IMCSE principles to enhance the user interface, data handling and analysis widgets. A demonstration case for interactive model-trading, including value, performance, and cost models with inherited data was completed to demonstrate impact on system decision making. Practitioner feedback validated the importance of model tradeoffs. Practitioner feedback validated the importance of model tradeoffs.

**4. Cognitive and Perceptual Considerations in Human-Model Interaction.** The research team investigated cognitive and perceptual considerations in human-model interaction. An analogy case on the transition from traditional to glass cockpits was performed to gather empirical knowledge on challenges and preliminary strategies for transitioning to complex interactive model-centric environments, toward the development of strategies to guide model developers and users.

---

## **RESEARCH RESULTS**

### The research results

- The researchers completed a review draft of the IMCSE research roadmap, which included specified research activities, interim outcomes and knowledge transfer actions. Two topic areas were explored in more depth, and findings were documented in conference papers for CSER 2016.
- The Interactive Epoch-Era Analysis framework and supporting tools were used to complete a defense-oriented demonstration case, focusing on opportunities to improve the uncertainty analysis, ease of use, data scaling, visualization techniques, and the overall analysis approach. Demonstration prototypes were developed for single epoch and multi-epoch analyses and were made available online. The research findings were documented in a conference paper for CSER 2016.
- A demonstration case for interactive model-trading, including value, performance, and cost models with inherited data was completed to demonstrate impact on system decision making. The research findings were documented in a conference paper for CSER 2016.
- The results of the investigation of the analogy case on the transition from traditional to glass cockpits were documented in a conference paper for CSER 2016. Preliminary cognitive challenges and mitigating heuristics were identified. Additional candidate analogy cases were identified for continuing investigation.

## **INTRODUCTION**

---

The IMCSE research program aims to develop transformative results through enabling intense human-model interaction, to rapidly conceive of systems and interact with models in order to make rapid trades to decide on what is most effective given present knowledge and future uncertainties, as well as what is practical given resources and constraints.

---

## **MOTIVATION**

Model based systems engineering (MBSE) methods and tools are used throughout the entire lifecycle to generate systems, software and hardware products, and work towards replacing documentation-based processes with more effective model-based methods. To take advantage of model-based techniques to develop systems, it is important to improve human and technology interaction to make trades and decide on what is most effective given the present knowledge and future uncertainties, as well as make logical decisions based on the availability of resources and constraints. The Interactive Model-Centric Systems Engineering (IMCSE) research program will inform and develop the SE methods, processes and tools to improve human-model interaction, with the goal of accelerating transition of SE to become a more model-based discipline.

---

## **INSUFFICIENCIES IN CURRENT PRACTICE**

Early concept decisions have always been critically important, and with continuously evolving systems of systems having long life spans, such decisions are now made throughout the entire life cycle. Soft factors become increasingly influential. For example, trust in model-based data sets and decisions are in part determined by the chosen model itself as perceived by specific decision makers. The timescale of making early architectural decisions is out of sync with the current model-based systems engineering capabilities and decision environments. New algorithms and novel modeling approaches must be discovered to accelerate technical and programmatic decision support from months to minutes. In order to effectively leverage and incorporate human knowledge and judgment, an interactive capability is needed. Much potential exists in maturing emerging novel methods for evaluating system responsiveness under complex uncertainties, to enable engineering of resilient systems. The IMCSE technical reports from prior phases highlight these. The Phase 3 Pathfinder Workshop Report (Rhodes and Ross 2015) discussed these insufficiencies, as identified by participants in the workshop.

## **PROJECT OVERVIEW: IMCSE**

---

Interactive Model-centric Systems Engineering (IMCSE), not to be confused with Model-based Systems Engineering (MBSE), is a research program that seeks to encourage the development of augmented complex systems thinking and analysis to support data-driven decision making

---

### **WHAT IS IMCSE?**

The IMCSE research program arises from the opportunity to investigate the various aspects of humans interacting with models and model-generated data. Future environments and practices need to leverage advancements in data science, visual analytics, and complex systems. IMCSE aims to develop transformative results through enabling intense human-model interaction, to rapidly conceive of systems and interact with models in order to make rapid trades to decide on what is most effective given present knowledge and future uncertainties, as well as what is practical given available resources and constraints. Open areas of inquiry include: how individuals interact with models; how multiple stakeholders interact using models and model generated information; facets of human interaction with visualizations and large data sets; and underlying fundamentals such as model purpose and model handling. IMCSE research is based on a belief that improving human-model interaction would significantly improve the utility of model-centric engineering. Human factor needs should be specifically considered, given models are an abstraction of reality and there are likely unique human factors considerations for understanding the representations. Developing the envisioned future for an interactive model-centric systems engineering experience will require new knowledge, new ways of working, and innovation in modeling constructs and technologies.

Systems scientists have long recognized that humans possess unique abilities for anticipation rather than simple reactive response. In order to increase the likelihood of developing complex systems that can deliver value to stakeholders across a dynamic, uncertain future, systems engineers must have both reactionary and anticipatory capacity to make better decisions. In contrast to reactionary capacity, which involves developing solutions after the fact, anticipatory capacity, as defined by Rhodes and Ross (2009), is “the capacity to continuously develop and apply knowledge acquired through a structured approach to anticipate 1) changing scenarios as stakeholder needs and systems context change over time; 2) to consider their consequences; and 3) to formulate design decisions in response<sup>1</sup>. Three key enablers of anticipatory capacity are mindset, methods, and environment. Models represent an abstraction of reality in order to make predictions about the future. Models can come in a variety of forms and formats, but fundamentally they are an encapsulation of reality that humans use to augment their ability to make sense of the world and anticipate future outcomes. Improvements in computation, simulation technologies, and human-machine interaction have created an opportunity to enable human-model interaction to greatly enhance anticipatory capacity. Complex, integrated models, of various levels of fidelity, can create large data sets in need of human pattern recognition skills.

---

<sup>1</sup> Rhodes, D.H. and Ross, A.M., "Anticipatory Capacity: Leveraging Model-Based Approaches to Design Systems for Dynamic Futures," 2nd Annual Conference on Model-based Systems, Haifa, Israel, March 2009.

Interaction enables real time interrogation of the data and opportunities for model creation as well as validation and learning. IMCSE is a research program intended to leverage human-model interaction in order to transform systems engineering decision making through anticipatory capacity.

---

## **RESEARCH PROGRAM VISION**

The vision for the IMCSE research program is to develop transformative results through enabling intense human-model interaction, to rapidly conceive of systems and interact with models in order to make rapid trades to decide on what is most effective given present knowledge and future uncertainties, as well as what is practical given resources and constraints. In order to accomplish this vision, IMCSE pursues a balanced basic and applied research approach. This draws on the strength of the academic environment (e.g. developing fundamentals, approaching with rigor, providing a neutral third party view of the problem). Additionally, IMCSE strives to keep the research relevant to the sponsor community, as well as enabling opportunities for knowledge and methods, processes, and tools (MPTs) transfer to sponsors. Such knowledge transfer opportunities include workshops, teleconferences and meetings, reports, papers, collaboration with other SERC activities, prototypes, methods, processes, and tools (MPTs), government partner applications, and potential student internships.

---

## **IMCSE PILLARS – FOUR TOPIC AREAS**

IMCSE is motivated by the convergence of four key topic areas: big data, visual analytics, complex systems, and model-based systems engineering. Each of these areas have associated with them large research and application efforts. This research program seeks to identify synergies and gaps at the intersection of these four topic areas, and leverage existing and new techniques in this area to create new knowledge and capabilities for systems engineering decision making. In order to focus the research program, early efforts are aimed to identify key challenges that summaries the gaps in the existing topic area overlaps.

---

### **BIG DATA**

We live in a world with big data. As data storage costs have shrunk, so too has the need for purging data. Additionally data is being generated through a large and growing number of means, from sensors to users to corporate IT environments. Even “document-based” data is becoming digital as technology (including OCR) becomes commonplace for capturing physical information as digital data. No consensus currently exists regarding a formal definition on what constitutes “big data,” but it is generally recognized as having a number of characteristics that make it “big.” One example description, from IBM<sup>2</sup>, characterizes big data as having challenges regarding Volume, Variety, Velocity, and Veracity. The challenge for Volume revolves around the scale of the data (e.g. how to store and recall large numbers of field entries in a database?). The challenge for Variety revolves around the different forms of data (e.g. how to store and compare data from photos, videos, blogs, articles, etc.?). The challenge for Velocity revolves around the

---

<sup>2</sup> [http://www.ibmbigdatahub.com/sites/default/files/infographic\\_file/4-Vs-of-big-data.jpg](http://www.ibmbigdatahub.com/sites/default/files/infographic_file/4-Vs-of-big-data.jpg)

analysis of streaming data (e.g. how to account for and parse large streams of potentially incomplete data in real time?). The challenge for Veracity revolves around the uncertainty of the data (e.g. different data sources have different degrees of trustfulness and reliability, so how to fuse data from such sources?).

The impact of big data is being felt across many fields from transportation to entertainment, education to banking, which will only increase as the benefit of leveraging such data becomes apparent. Such benefits have been recognized by a growing number of commercial organizations who are leveraging this inundation of data to gain insights into phenomena to create predictive models (e.g. of user behavior and preferences). For example, Amazon and Netflix both have sophisticated user preference models that are used to make recommendations to users based on their own (and related others) browsing and shopping/viewing history. Additionally, Netflix has used this information (and Amazon recently as well) to generate design requirements for new shows. House of Cards, produced by Netflix, was partially designed based on derived preferences of its viewer base in order to increase the perceived value of the program<sup>3</sup>.

While not necessarily generated in a similar manner, DoD has already a vast amount of data stored in documents, for example requirements documents, design documents, DoDAF, etc, which represent latent data that could be leveraged using techniques being developed in the commercial application space. What would a ground vehicle recommendation look like? How would it parse and analyze historical requirements documents and contextual information in order to predict and/or augment modern user needs? Big data is a topic area that holds promise in providing a foundation for large scale analytics to predict the future.

---

#### **VISUAL ANALYTICS**

*Visual analytics is resulting in a transformative capability, bridging human and computer analysis.*

Visual analytics is a topic area that has likewise been a growing area for research and application. At its core, visual analytics is about collaboration between human and computer using visualization, data analytics, and human-in-the-loop interaction. More than just visualization tools, visual analytics aims to take advantage of a human's ability to discover patterns and drive inquiry in order to make sense of data. In 2007, DHS sponsored the National Visualization and Analytics Center, which developed a research agenda called Illuminating the Path. In it, visual analytics was defined as "the science of analytical reasoning facilitated by interactive visual interfaces" that "provides the last 12 inches between the masses of information and the human mind to make decisions."<sup>4</sup> Application areas range from homeland security to anti-fraud, banking to insurance. One common element in much of the current visual analytics work involves case applications comparing VA-supported inquiry results to ground truth, that is, discovery of patterns in "natural" data. One consequence of these studies is that the validity of the

---

<sup>3</sup>[http://www.nytimes.com/2013/02/25/business/media/for-house-of-cards-using-big-data-to-guarantee-its-popularity.html?pagewanted=all&\\_r=0](http://www.nytimes.com/2013/02/25/business/media/for-house-of-cards-using-big-data-to-guarantee-its-popularity.html?pagewanted=all&_r=0)

<sup>4</sup> Jim Thomas, Director, USDHS National Visualization and Analytics Center, "Visual Analytics: An Agenda in Response to DHS Mission Needs," 2007

applications can be compared to observable “truth.” This allows researchers to test how well their predictive models match reality, for example, using VA to discover hackers trying to break into streams of ATM data; or discovering patterns of use in bike sharing programs as a function of time and geography. In both of these examples there are “real” processes at play and actual measurable real world data against which to validate predictions by the human-machine VA system. VA has been shown to be incredibly useful for developing models of natural data.

---

## COMPLEX SYSTEMS

*Developing complex systems necessitates an approach to generate, manage, and analyze artificial data across all aspects of system complexity.*

Our application domain is the development of (artificial) systems that serve the purpose of delivering value to stakeholders. By “artificial” we mean that these systems are artifacts created by humans for a purpose, to be contrasted with natural systems, which are not created by humans. Over time, the complexity of systems has tended to grow, not only due to scale and interconnectedness, but also due to increased scope in our ability to describe the system. This enhanced scope reflects realization that the success of artificial systems requires a fuller understanding of how the system is structured, behaves, performs in different contexts, performs over time, is perceived across stakeholders<sup>5</sup> <sup>6</sup>. This means that to describe a complex system, one must consider all five perspectives, thereby creating a richer description of the system. Developing complex systems necessitates an approach to generate, manage, and analyze artificial data across these five aspects.

The growing complexity of systems is well-recognized, and investigation of system complexity as related to engineered systems is an active subject of inquiry. Complexity, for instance, can relate to the number of constituent and component interconnections, and to the necessary rapid rate of information generation and exchange. It can also relate to emergent behavior as a result of interactions of constituent systems in a system of systems.

**Defining Systems Complexity.** Many authors have and continue to define system complexity. Gasper (2012)<sup>7</sup> discusses three bodies of work than can be used as a basis for complexity definition in the context of engineering. Herbert Simon (1962)<sup>8</sup> proposes that how complex or simple a structure is depends critically on the way in which we describe it. Simon proposes a hierarchical approach to complexity, decomposing the system until it can be understood.

---

<sup>5</sup> Rhodes, D.H., and Ross, A.M., "Five Aspects of Engineering Complex Systems: Emerging Constructs and Methods," 4th Annual IEEE Systems Conference, San Diego, CA, April 2010

<sup>6</sup> Gaspar, H., Rhodes, D.H., Ross, A.M., and Erikstad, E.O., "Addressing Complexity Aspects in Conceptual Ship Design: A Systems Engineering Approach" *Journal of Ship Production and Design*, Vol. 28, No. 4, Nov 2012, pp. 145-159.

<sup>7</sup> Gasper, H., Rhodes, D., Ross, A. and Erikstad, E., "Addressing complexity aspects in conceptual ship design: a systems engineering approach", *Journal of Ship Production and Design*, Vol. 28, No. 4, November 2012, pp. 1–15

<sup>8</sup> Simon, H., "The architecture of complexity", *Proceedings of the American Philosophical Society*, 106, 6, 467–482, 1962

Kolmogorov (1983)<sup>9</sup> definition of complexity asserts the more information an object has, the more complex it is. Given the system is the object, complexity can be understood as related to the other objects that interact with the system. The specification of an object is easier when another object to which this object has a relation is already specified. A third work by Suh (2005)<sup>10</sup> discusses the idea of information connected to the design complexity, proposing that the violation of the information axiom, to minimize the information content of the design will maximize the probability of success, will result in complexity in the system.

**Types of System Complexity.** Structure and behavior are the two aspects of complex systems addressed in classical model-based systems engineering<sup>11</sup>. Rhodes and Ross (2010)<sup>12</sup> propose five essential aspects for the engineering of complex systems: structural, behavioral, contextual, temporal, and perceptual. They argue that the contextual, temporal and perceptual aspects have been under-addressed in engineering methods, and have past and ongoing research efforts on advancing the constructs and methods for contextual, temporal, and perceptual aspects. Response Systems Comparison is a resulting method to address the five aspects<sup>13</sup>. The method has been applied in various domains and for various types of problems, for example, Gasper<sup>7</sup> describes the application for a conceptual ship design problem.

---

<sup>9</sup> Kolmogorov, A. N., "Combinatorial foundations of information theory and the calculus of probabilities, Russian Mathematical Surveys, 38, 4, 27–36, 1983.

<sup>10</sup> Suh, N. P., "Complexity—theory and applications", Oxford University Press, Oxford, UK, 2005

<sup>11</sup> Oliver, D., Kelliher, T., Keegan, J., *Engineering Complex Systems with Models and Objects*, NY: McGraw-Hill, 1997.

<sup>12</sup> Rhodes, D.H. and Ross, A.M., "Five aspects of engineering complex systems: emerging constructs and methods", Proceedings, 4th Annual IEEE Systems Conference, April, San Diego, CA, 2010

<sup>13</sup> Ross, A.M., McManus, H.L., Rhodes, D.H., Hastings, D.E., and Long, A.M., "Responsive Systems Comparison Method: Dynamic Insights into Designing a Satellite Radar System," AIAA Space 2009, Pasadena, CA, Sep 2009

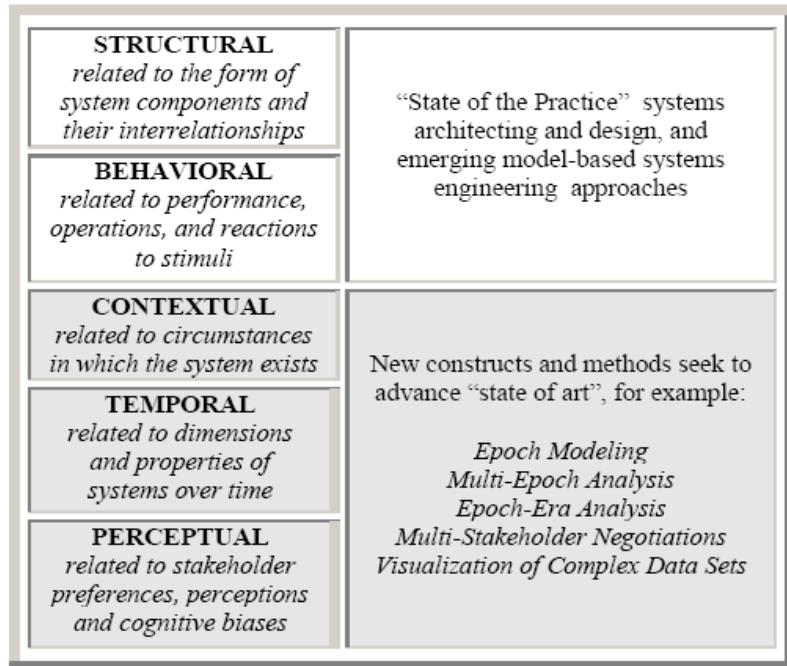


Figure 1. Five Aspects of Complex Systems<sup>12</sup>

**Human-System Interaction Complexity.** The complexity of the human-system interaction considerations is increasingly important in developing complex systems. A 2007 report of The National Academies<sup>14</sup> presents a discussion of the challenges, with research and policy recommendations. Many of the points brought out in this report and in subsequent work extend to understanding of complex systems. A number of the recommendations have extensions to the challenges of we see for IMCSE, and are beginning to be addressed through research. Examples include:

- Remote collaboration is difficult to participate in or observe without proper remote collaboration tools enabling interactivity of human to human, and human to model.
- Cognitive and perceptual limitations constrain the amount of information that can be considered at a point in time by a single decision maker; multi-sensory representations may allow for some loosening of this constraint and improve human-model interaction.
- Research has increasingly uncovered the important role of context effects on both systems in use, design, and on the decision makers themselves. Facilities that can represent and control for these context effects may uncover approaches for mitigating or taking advantage of these effects.

---

<sup>14</sup> National Research Council (2007), *Human-Systems Integration in the System Development Process: A New Look*, Committee on Human-System Design Support for Changing Technology, R.W. Pew and A.S. Mavor, Eds., Committee on Human Factors, Division of Behavioral and Social Sciences and Education, Washington, DC: The National Academies Press.

---

## **MODEL-BASED SYSTEMS ENGINEERING (MBSE)**

*Model-based systems engineering generates “artificial data” about our systems which we use to make decisions that impact the future/continuing success of that system*

Traditional systems engineering has been document-heavy and process-driven, resulting in many opportunities for miscommunication and mistakes during “hand-offs” between phases and teams. Models are often used during design and development in order to predict behavior or other consequences of design decisions, before the system is built or operated. In contrast to document-based engineering, “model-based systems engineering (MBSE) is the formalized application of modeling to support system requirements, design, analysis, verification, and validation activities beginning in the conceptual design phase and continuing throughout development and later life cycle phases.”<sup>15</sup> Today, however, standalone models are typically related through documents.

A future vision is for organizations to use “shared system model(s) with multiple views, and connected to discipline models,” in order to reduce effort creating and aligning documents, and to increase synthesis and coherence across disciplines throughout design<sup>16</sup>. Regardless of the degree to which MBSE is employed, its benefits stem from moving to models to represent systems with less ambiguity, more parsimony, and more consistency, resulting in reduced acquisition time, enhanced reliability, etc. MBSE generates “artificial data” about systems which can be used to make decisions that impact the future and continuing success of that system.

---

## **CHALLENGES ACROSS THE PILLARS**

Each of the four topic areas above are themselves large areas of active research and development across government, academia, and industry. IMCSE in particular is interested in the intersection of these four areas with application to improving systems engineering decision making (Figure 2).

More than just applied visual analytics, IMCSE seeks to look at data generated by models, in order to make better decisions in how to deliver sustained value to stakeholders. IMCSE research activities are addressing significant challenges, including:

- 1) Visual analytics of artificial (i.e. model-generated) data: how does this differ from VA of natural data? How to take into account the impact of various model implementations on pattern finding and matching of mental and constructed models? How to validate predictions without ground truth available?
- 2) Active tradeoffs of models themselves: too often models are used without sufficient investigation into the impact of the models on the data being used for decisions; these include performance models, cost models, and value models. Model selection fundamentally impacts the patterns to be discovered in the artificial data.

---

<sup>15</sup> INCOSE SE Vision 2020 (INCOSE-TP-2004-004-02, Sep 2007)

<sup>16</sup> [http://www.omgwiki.org/MBSE/doku.php?id=mbse:incose\\_mbse\\_iw\\_2014](http://www.omgwiki.org/MBSE/doku.php?id=mbse:incose_mbse_iw_2014)

- Big Data + Visual Analytics...  
+ Complex Systems + MBSE = IMCSE
  - Volume, variety, velocity, and veracity of data
  - Collect data, visualize, interact, model, find patterns, generate insights, repeat
  - Structural, behavioral, contextual, temporal, and perceptual complexities
  - Integrated models including requirements, structure, behavior, parametrics
- Potential use for this merged capability for decision support within and across systems engineering throughout lifecycle

On the power of humans with computers:

*"statistics (computing) + humans is much more powerful than statistics alone or humans alone"*

– Professor Remco Chang, Tufts University  
Visual Analytics Lab, Aug 2013



Interactive Model-Centric Systems Engineering	MBSE	STRUCTURAL	form of system components and their interrelationships	Existing "state of practice" systems architecting and model-based systems engineering
		BEHAVIORAL	function/performance, operations, and reactions to stimuli	Emerging "state of art"
		CONTEXTUAL	circumstances in which the system or enterprise exists	<i>Epoch Modeling Multi-Epoch Analysis Epoch-Era Analysis Dynamic Tradespace Exploration Multi-Stakeholder Negotiations Comprehension of Complex Datasets Studies of Decision Making and more...</i>
		TEMPORAL	dimensions and properties of systems over time	
		PERCEPTUAL	stakeholder preferences, perceptions and cognitive biases	

Developing complex systems necessitates an approach to generate, manage, and analyze artificial data across these five aspects, which result in improved SE decision making

Figure 2. Research at the Intersection of four pillars

## **RESEARCH PROGRESS**

---

The Phase 3 research focused on five activities. The first was the development of an initial research roadmap extending from the prior phase workshop, serving as a review draft to elicit inputs from the broader systems community. In support of the roadmap development, the research team investigated two research themes that emerged from the pathfinder investigation in Phase 2, including non-technical challenges for digital system models and model curation. A third research activity was the investigation of cognitive and perceptual considerations for human-model interaction, including an analogy case study. Continuing from the prior phase, the research team evolved the Interactive Epoch-Era Analysis framework and demonstrated the application of the framework in a demonstration case. The fifth research activity continued the model choice and tradeoff investigation, with demonstration cases.

---

## **IMCSE RESEARCH ROADMAP**

IMCSE research is complementary to other model-based research and strategic initiatives. This roadmap does not duplicate model-based engineering topics, but focuses specifically on research unique to the interactions of humans with and through models and model-generated information, as well as stewardship of model-centric environments.

---

## **DEVELOPING THE RESEARCH ROADMAP**

The IMCSE pathfinder project in Phase 1 and 2, along with current phase efforts included face-to-face gatherings of stakeholders to create a research agenda. An initial workshop held during Phase 2 seeded the initial research agenda<sup>17</sup>.

There have been many contributors to the roadmap though multiple activities including the 2015 IMCSE Pathfinder Workshop, technical workshops and meetings, graduate student research, and interchanges with SERC leadership and collaborating researchers. The current version of the roadmap is published as a review draft to elicit additional feedback and contributions.

As shown in Figure 3, the roadmap is comprised of four focus areas (1) guided human-model interaction; (2) multi-stakeholder interaction within model-centric environments; (3) enabling model-based decisions; and (4) curation of model-centric environments. The roadmap is included as Appendix B of this report. The figure shows the enterprise, which contains one or more interactive model-centric environments. A key area of the roadmap covers the curation of the model centric environment, which is comprised of the model-based assets and practices, and the people how interact with models in various ways. Research is needed to understand the various use cases for guided human-model interaction, for different types of model developers and users. Frameworks and technologies are needed to enable model-informed decisions by the various

---

<sup>17</sup> D. Rhodes, A. Ross. Interactive Model-Centric Systems Engineering (IMCSE) Pathfinder Workshop Report, Systems Engineering Research Center, Systems Engineering Advancement Research Initiative; January 2015.

types of stakeholders within the enterprise. Additionally, multi-stakeholder interaction through the power of model-centric environments must be effectively enabled.

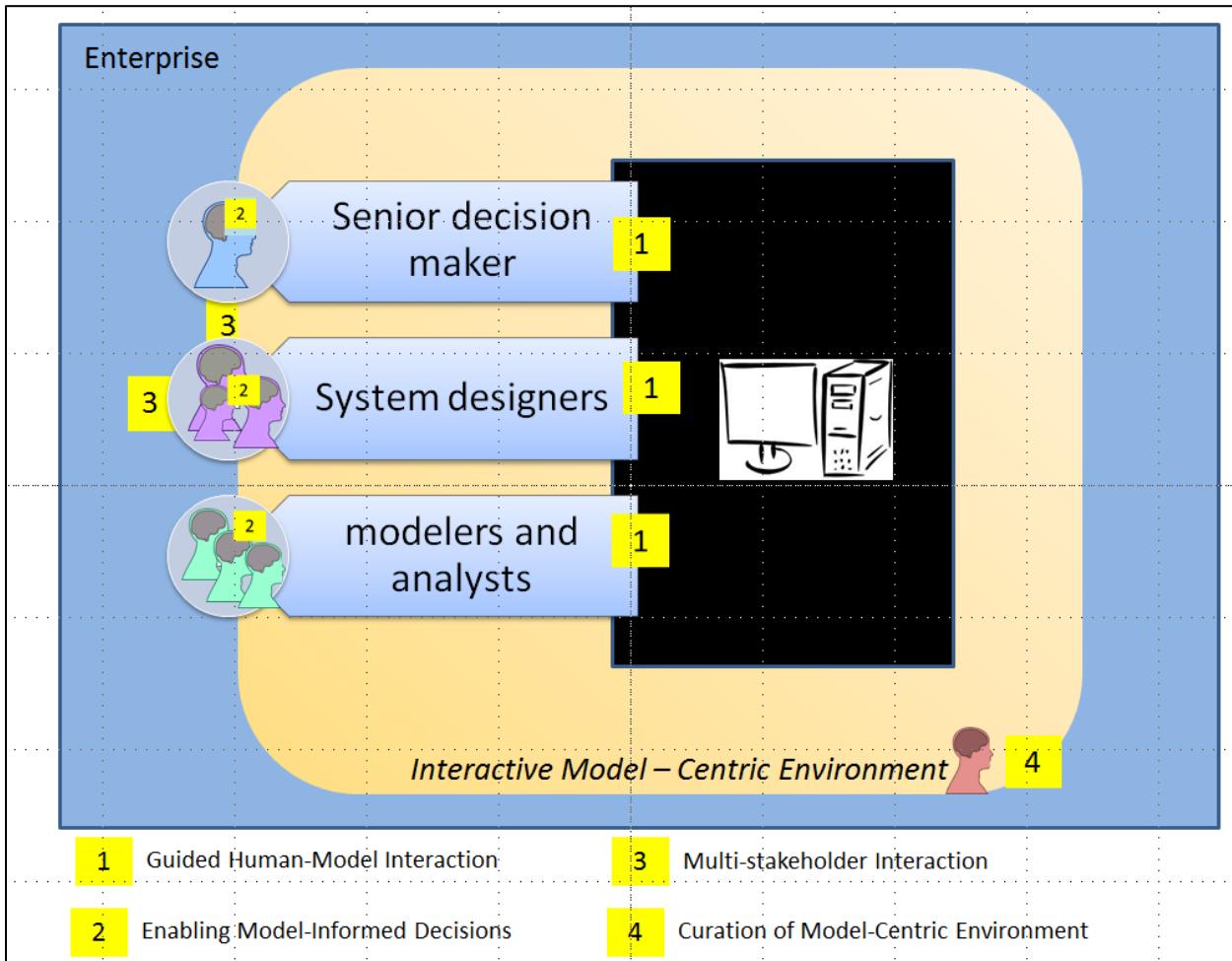


Figure 3. Four Focus Areas of IMCSE Research Roadmap

Given the footprint of IMCSE, it would not be possible to convene a large enough community in a participant workshop for the purpose of a collaboratively-derived research agenda. The research team explored various approaches that have been used, and identified a strategy for ongoing work. The goal is to be able to engage a large and diverse community around the research agenda, and determine an approach that may include both face-to-face and virtual activities.

---

## INVESTIGATION OF SELECTED RESEARCH THEMES

As the research team developed the roadmap, two themes were further investigated to inform the roadmap development (in addition to other phase 3 tasks). These were non-technical challenges of digital system models and an investigation of the need for a model curation function in model-centric environments.

---

### DIGITAL SYSTEM MODEL<sup>18</sup>

Models are increasingly used to drive major acquisition and design decisions, yet the diverse set of model developers, analysts, and decision makers are faced with many challenges in transitioning to a model-centric paradigm. While the technical challenges are relatively well understood and there is significant progress being made in addressing them, the non-technical aspects have not received the same level of attention. During this phase, the research team performed a preliminary investigation of the non-technical challenges for the *Digital System Model* in context of the larger topic of the digital framework including Digital Thread and Digital Twin.

#### Introduction

The US Department of Defense (DoD) envisions a model-centric environment, comprised of the Digital Thread, Digital Twin, and Digital System Model (collectively referred to in this report as a digital framework), in which it and its industry partners can jointly engage in model-based engineering. Under this digital framework, information will be shared rapidly and effectively shared, systems will be designed holistically without losing detailed information, and the entire acquisition process may be greatly accelerated.

While clear, though surmountable, technical challenges exist in the realization of this digital framework, such as computing power and model fidelity, this investigation focused on non-technical challenges. These have not been adequately examined in previous literature on the topic, particularly issues of intellectual property (IP) and knowledge assessment (KA). IP in this context refers to any piece of information owned or held in some way by a person or organization. In the defense industry, this includes patents, software (or algorithms more generally), material data, business approaches, trade secrets, etc. A key characteristic of IP is that, while information can be easily shared, value exists in information asymmetries (i.e. possessing information or IP that others do not). Since the construction of such a digital framework fundamentally requires the integration of IP from multiple sources, issues of valuation, compensation, and protection of IP arise. KA refers to the assignment of validity to any particular piece of information or expertise. This can be a technical process, such as double-checking statistical methods used to evaluate the results of an experiment, or a more social process, such as deciding which news source to trust. KA issues include “buy-in” (i.e. trust) both by technical staff directly creating the models and by decision-makers in the acquisition and operation processes, as well as inter-model validity assessment in the cases of multiple models

---

<sup>18</sup> Reid, J.B. and Rhodes, D.H., “Digital System Models: An investigation of the non-technical challenges and research needs”, Conference on Systems Engineering Research, March 2016.

operating in the same domain (e.g. two crack-growth modeling methods). This report uses the following definitions:

Digital Thread. The Digital Thread (DTh) is analytic framework for integrating technical data, costs, predictions, and other accumulated knowledge over the course of design, development, and operation of a system. It is intended to provide ready access to usable information for decision makers during the design process. It includes tools such as tradespace analysis and visualization tools. The term “digital thread” is sometimes referred to by other terms, for example, Lockheed Martin Corporation has used the term Digital Tapestry<sup>19</sup>. The DoD defines Digital Thread as: “An extensible, configurable and component enterprise-level analytical framework that seamlessly expedites the controlled interplay of authoritative technical data, software, information, and knowledge in the enterprise data-information-knowledge systems, based on the Digital System Model template, to inform decision makers throughout a system's life cycle by providing the capability to access, integrate and transform disparate data into actionable information”<sup>20</sup>.

Digital Twin. A Digital Twin (DTw) is an integrated model of an as-built system including physics, fatigue, lifecycle, sensor information, performance simulations, etc. It is intended to reflect all manufacturing defects and be continually updated to include wear-and-tear sustained while in use. The goal is to more effectively and safely manage the individual product as well as to facilitate investigations of potential design or operational changes on the health of the system. The DoD defines Digital Twin more specifically as: “An integrated multiphysics, multiscale, probabilistic simulation of an as-built system, enabled by Digital Thread, that uses the best available models, sensor information, and input data to mirror and predict activities/performance over the life of its corresponding physical twin”<sup>20</sup>.

Digital System Model. The Digital System Model (DSM) is essentially the proposed product of MBE<sup>21</sup>. It is the integrated model of all technical data out of which individual Digital Twins will be constructed, and is the technical grounding that the decision-making analytics of the Digital Thread refers to. The DoD defines DSM as: “A digital representation of a defense system, generated by all stakeholders that integrates the authoritative technical data and associated artifacts which define all aspects of the system for the specific activities throughout the system lifecycle.<sup>20</sup>”

The digital framework, as comprised of DTh, DTw and DSM, has many potential benefits for the DoD acquisition process, including:

- More rapid and cheaper cycling of design concepts in a form of virtual prototyping<sup>21</sup>

---

<sup>19</sup> *Digital Tapestry*, Lockheed Martin. Available: <http://www.lockheedmartin.com/us/what-we-do/emerging/advanced-manufacturing/digital-tapestry.html> [Accessed: 10/15/2015]

<sup>20</sup> *Glossary: Defense Acquisition Acronyms and Terms*. Department of Defense, US Department of Defense, Defense Acquisition University, Learning Capabilities Integration Center, Center for Acquisition and Program Management, 2015

<sup>21</sup> P. Zimmerman. MBSE in the Department of Defense. Office of the Deputy Assistant Secretary of Defense for Systems Engineering, DoD, 2015

- Better analysis and comprehension of complicated systems<sup>21</sup>
- Increased reuse of technical data and models from one project to another, as well as from one stage of a project to another stage of the same project (reduces stove-piping)<sup>22</sup>
- Reversal of the trend of outsourcing technical knowledge from the DoD to contractors and industry partners<sup>23</sup>

DoD and various private firms have actively begun developing and implementing the digital framework or portions of it. For example, the DoD's Computational Research and Engineering Acquisition Tools and Environments (CREATE) program has developed several modelling tools for the design phase<sup>24</sup>. Lockheed Martin used the same 3D solid models of the F-35 Lightning II for engineering design, manufacturing, tooling, and development of training and maintenance materials<sup>25</sup>. The United States Air Force (USAF) is currently testing elements of digital framework on four different acquisition programs<sup>23</sup>. NASA is working on developing Digital Twins for a microelectromechanical system<sup>26</sup>.

Non-defense applications of the digital framework exist as well. GE is developing Digital Twins of wind farms<sup>27</sup> and Singapore's partnership with Dassault Systèmes is creating a Digital Twin of the city of Singapore as a whole<sup>28</sup>. The interest by other domains highlights expected usefulness of digital models, and the assumption that technical challenges posed in their development are believed to be surmountable.

### **Non-Technical Challenges**

The technical challenges involved in creation of DSM are better understood (many since before DSM itself existed as a distinct concept<sup>29</sup>) and are currently being worked on by various parties in government<sup>30</sup>, industry<sup>19,31</sup>, and academia<sup>32</sup>. While IP issues have occasionally been acknowledged in discussions of DSM<sup>29,32</sup>, there is little evidence these have been addressed (with the exception of cybersecurity). Similarly, while KA challenges have been examined by some researchers<sup>17</sup>, they have not been considered as priorities. These non-technical challenges should

---

<sup>22</sup> P. Zimmerman. A Framework for Developing a Digital System Model Taxonomy. *18th Annual NDIA Systems Engineering Conference*; 2015

<sup>23</sup> G. Warwick. USAF Selects Lead Programs for 'Digital Twin' Initiative. *Aviation Week & Space Technology*; 2015

<sup>24</sup> E. M. Kraft. HPCMP CREATE TM-AV and the Air Force Digital Thread. *SciTech*; 2015.

<sup>25</sup> The Digital Thread Key to F35 Joint Strike Fighter Affordability, Lockheed Martin. Available: <https://www.onlineamd.com/article/amd-080910-f-35-joint-strike-fighter-digital-thread> [Accessed: 9/22/16].

<sup>26</sup> E. Glaessgen, D. Stargel. The Digital Twin paradigm for future NASA and US Air Force vehicles. 53rd Struct. Dyn. Mater. Conf.; 2012 1-14.

<sup>27</sup> T. Kellner. Wind in the Cloud? How the Digital Wind Farm Will Make Wind Power 20 Percent More Efficient. *GE Reports*; 2015

<sup>28</sup> G. Goh. Building Singapore's 'digital twin'. *Digital News Asia*; 2015

<sup>29</sup> *Final Report of Model Based Engineering (MBE) Subcommittee*, NDIA Systems Engineering Division M&S Committee, National Defense Industrial Association2011

<sup>30</sup> A. Feeney. *Digital Thread for Smart Manufacturing*. Engineering Laboratory-Systems Integration Division-Systems Engineering Group, US National Institute of Standards and Technology, 2014.

<sup>31</sup> M. A. Aisenberg, D. Davis, and T. D. Kehoe. Intellectual Property Rights for Digital Design and Manufacturing: Issues and Recommendations, The MITRE Corporation, 2014.

<sup>32</sup> T. D. West, A. Pyster. Untangling the Digital Thread: The Challenge and Promise of Model-Based Engineering in Defense Acquisition. *INSIGHT* 2015; vol. 18, 45-55.

not be downplayed or neglected. While overcoming the technical challenges addresses DSM power and capability, if left unaddressed, IP and KA issues could substantially reduce the effectiveness of DSM. Working to identify and address such issues may proactively prevent DSM from the risk of becoming an expensive, powerful, but inadequately used tool in DoD's acquisition and operations portfolio.

The DSM itself is highly valuable IP. As the integration of all "authoritative data and associated artifacts" possession of the DSM could allow for full understanding of the system and its individual subsystems. Many aspects of manufacture and assembly would likely be either directly included in or could be indirectly inferred from the DSM. The simulation and design assessment capabilities of DSM and the digital framework as a whole significantly enhance the ease of understanding the system. These aspects are an integral part of the potential benefits of DSM, yet also a critical area of concern for cybersecurity. Where previously, an attacker would have to steal hundreds or thousands of documents from dozens of organizations and then sort through these documents in order to understand the system. With DSM, the attacker could have a single target to go after, a target that is designed for easy learning. Additionally, since the intention of DSM is for the same model to be used throughout the acquisition process, an attacker could potentially introduce malicious changes into the model during the manufacture stage that would become instantiated in the actual product. Cybersecurity is not a novel concern to the DoD, however, and the cybersecurity implications of DSM have been well understood for several years<sup>11,33</sup>.

Beyond cybersecurity, the DSM must have buy-in from its various collaborators. KA issues will be discussed in a later section (Section 0) but a relevant key issue here is the sharing of IP among competitive firms. If one firm owns the system being designed, from start to finish, then DSM would be a relatively simple matter concerning IP. If the same firm also operated the product, then the development of the Digital Twin from the DSM would likewise be a straightforward IP matter. This is not typically the case in the defense realm. The DoD sets requirements and is the operator of the end product. Multiple teams of firms then bid on the project. Even after one team has been selected, this team will often use other firms as suppliers. These teams are not fixed groups; on one project two defense contractors may be cooperating, while on another simultaneous project they may be competitive rivals. As a result, these firms may be unwilling to share the detailed IP required for the DSM. One approach to addressing this issue in the past has been through a combination of non-disclosure agreements (NDAs) and a design process where subsystem interfaces are rigidly defined but the actual subsystems themselves are largely treated as black boxes<sup>34</sup>. This allows one firm to maintain control over their IP, while delivering only the final product to the larger working group. Overcoming the design limitations that this approach introduces is a major goal of DSM, which seeks to more holistically model the system, and accordingly new methods of addressing IP concerns will be required.

Patents. Patents are a potential tool for safeguarding the interests of DSM stakeholders. Jointly developed technologies for simulation and model integration methods could be protected under

---

<sup>33</sup> M. McGrath. Protecting the Digital Thread. *Global Supply Chain Summit*; 2014

<sup>34</sup> *Systems Engineering Guide for Systems of Systems*. Systems and Software Engineering, Department of Defense, 2008

patents shared by the developers. Under the Cooperative Research Act and its amendment, the Standards Development Organization Advancement Act, cooperative research towards developing a DSM standard would potentially be protected from antitrust litigation, as it is a research area of national interest<sup>35</sup>. This is a ready extension of activities that the Digital Manufacturing and Design Innovation Institute (DMDII) is already pursuing<sup>31</sup>.

It should be noted that the use cases of patents in the defense industry are somewhat constrained in this context. First, all patents must be publically published by the US Patent and Trademark Office (USPTO), which is clearly contrary to the classified nature of many defense technologies. Second, the cooperative development style described above only works for shared technologies, which applies primarily to the development of the DSM structure. For the systems and subsystems being modeled, many other technologies will be privately held by an individual firm. While some of these technologies could be protected under patent law, many valuable forms of IP, such as material properties, are either not patentable or not worth being patented. Overall, patents will be useful in the process of developing the enabling structure of the DSM in general (such as the modeling packages, the data exchange methods, and the physical computational and storage devices that constitute DSM), but less so in protecting the content of specific applications of DSM (such as fuselage designs, experimental data on a new composite material, or integrated circuit design).

**Copyright.** Copyright has limited application in the protection of DSM. While specific presentation of facts, such as experimental results or material properties, are subject to copyright, the facts themselves are not, and thus a restatement in a different format is enough to circumvent copyright. As a result, copyright in general is not relied upon for protecting such information. Similarly, while computer programs, such as modeling software, are copyrightable, the fundamental basis of these (e.g., specific mathematical methods) are not. This results in variation in the usefulness of copyright as it pertains to software. In cases where the specific software package is itself valuable (e.g., Microsoft Word) copyright is a sufficient tool of protection. In other cases where the value is not in the software package but in the technical underpinnings, such as a proprietary modeling program, copyright may not sufficiently guard the IP. Furthermore, while copyright might protect the modeling software itself, it would not protect the models made of products, which are valuable in their own right.

**Trade Secrets.** Currently, technology and methods not patented or copyrighted are typically protected via trade secrets. While trade secrets enjoy some legal protections<sup>36</sup>, these protections are contingent on the IP being kept secret. Unless the DoD is the sole holder of the DSM, in order for the DSM to effectively operate, information will need to be shared among firms. One of the more common methods for handling inter-firm sharing of trade secrets is the NDA. While useful, NDAs have their drawbacks when it comes to DSM. For one, they are time-consuming to complete, which could be an issue if they need to be continually updated during an iterative design process. Additionally, the level of detail of information being shared under

---

<sup>35</sup> Standards Development Organization and Advancement Act of 2004, U. S. Congress, 2004

<sup>36</sup> Trade Secret, WEX. Legal Information Institute, Cornell Law School, Ithaca, NY.

DSM is quite high. Firms may not feel that NDAs are substantial enough to protect this information, particularly as the difficulties in discovering violations of NDAs might be difficult or impossible, limiting enforceability of such contracts<sup>21</sup>. Excessive use of NDAs could limit use of knowledge generated by one project in another project, one of the proposed goals of DSM.

**Comparisons.** While some IP issues are unique to DSM, many are not. Other industries and projects that have dealt with these problems may provide insight through either similarity or contrast. In this section, two such comparisons are briefly described.

*Standards Organizations.* Standards organizations exist in many industries. Some are private groups, such as W3C or IEEE, while others are governmental, such as ANSI or ISO. Depending on the group and the specific standard, compliance may be either voluntary or mandatory, but either way, a common reason for the existence of standards is interoperability of products between firms. While other reasons for standards exist, including health and safety, these are less relevant to the topic of DSM. In general, the development of interoperability standards is motivated by the firms themselves, seeking to eliminate competition in certain dimensions (such as size of customer-base and location of market) and facilitate competition based on the products themselves. In order to avoid violation of antitrust laws, these standards are typically open, meaning that any firm may develop products under the standard. If there is protected IP, such as patents involved in the complying with a standard, then this IP must be licensed under fair, reasonable, and non-discriminatory (FRAND) terms<sup>37</sup>. Additionally, these standards are typically formalized through a consensus method, though the exact definition of consensus varies from one standards organization to another.

In the DoD context, there is little incentive for the industry firms to develop specialized DSM interoperability standards as the size of their customer base is fixed at one, the DoD. Thus, interoperability will not allow an expansion of market. The firms may develop their own, limited, proprietary version of DSM in order to develop better products and thus better compete<sup>25</sup>, but they will not likely reach out to other firms without external incentives. The power held by being the sole customer, though, does allow the DoD significant leeway in providing those external incentives. If the DoD defines specific interoperability requirements of modeling software or even requires specific modeling software to use, these requirements will be followed by the industry.

It should be noted that as the DoD would be the motivating force behind the standardization effort, the DoD is not bound by any consensus method of DSM development, thought it may wish to do so in order to make use of industry expertise and in general to promote goodwill with the industry. Additionally, due to the aforementioned exemptions in antitrust laws (Section 0), it may be possible that the IP used to develop DSM does not need to be licensed under FRAND terms.

---

<sup>37</sup> R. Stim, F. Stephan. *What To Do If An NDA Is Violated*, NDAS For Free, Available: [www.ndasforfree.com/4StepstoTake.html](http://www.ndasforfree.com/4StepstoTake.html) [Accessed: 2/2/16].

This would mean additional options are open to the DoD regarding the enabling structure for the DSM.

Federal Drug Administration's Sentinel Initiative. Beginning in 2008, the Federal Drug Administration (FDA) developed the Sentinel Initiative, a new method of active surveillance (in this context surveillance refers to continued monitoring of the efficacy and side effects of medical products post-approval)<sup>38</sup>. Instead of primarily relying upon passive surveillance (e.g. unsolicited reports from healthcare providers, from consumers, and from medical product manufacturers), the FDA sought the capability to actively query medical record data sets to answer safety questions. While medical records are not IP in the same sense as what has been discussed previously, they are considered protected information under such laws as the Health Insurance Portability and Accountability Act (HIPAA). As a result the current holders of the data sets, including healthcare providers (such as hospitals) and health insurance firms, are legally unable to provide these data sets to other parties, including the FDA or academic researchers. The Sentinel Initiative circumvents this restriction by operating on a distributed system, where data sets never leave their current location. The FDA sends out a specific query to all data partners (possessors of medical records who are part of the Sentinel Initiative). This query is then processed by each data partner separately, the results returned in aggregate to the FDA without identifying information. In this way the FDA can assess whether drug X is causing side effect Y without violating the privacy of any patients.

In the DSM context, while firms are not legally required to guard their IP, their natural unwillingness to share valuable information accomplishes a similar effect to the HIPAA restrictions. In this way, the Sentinel Initiative suggests a potential workaround to this problem. It is possible that, from a firm-to-firm perspective, the modeling software and the models generated with that software may be treated as black boxes with each firm only have access to inputs and outputs of the simulations of the others.

### **Knowledge Assessment**

Knowledge assessment is defined as the assignment of validity to any particular piece of information or expertise.

**Generating Buy-In.** Buy-in can be defined as a combination of trust and willingness in a tool, as well as the ability to use it. This is a non-trivial matter. It is all too easy in any organization for some newly developed tool to sit unused and unmaintained if those who should use it lack the skillset, willingness, or trust for it. Lack of buy-in to a new initiative can a significant and detrimental impact on the enterprise as a whole but can be avoided<sup>39</sup>.

Developing the appropriate skillset requires investment in training, reference materials, and experience. Trust and willingness are not as straightforward to achieve. On the technical side, establishing trust requires details, examples, and authority. This means that DSM users should

---

<sup>38</sup> *The Sentinel Initiative*. Office of Surveillance and Epidemiology, Federal Drug Administration, 2010

<sup>39</sup> K. Thomson, L. de Chernatony, L. Arganbright, and S. Khan. The Buy-in Benchmark: How Staff Understanding and Commitment Impact Brand and Business Performance. *Journal of Marketing Management* 1999; vol. 15, 819-835.

have access to data on technical underpinnings and assumptions for the various models, as well as proof of validation in the form of examples (these can also serve as useful reference materials for developing expertise). This does not necessarily need to be immediately presented to every user at all times, but should be accessible by those who desire it, and at varying levels of detail for different use-cases. Some of this information can be elided in certain cases by referencing standard validation certifications, though these certifications must themselves be trusted by the user and are thus unlikely to be available during the initial implementation of a new tool. DoD has an organizational standard in place useful for cataloging this information and making it accessible to users: DoD Metadata Registry and the Modeling & Simulation Catalog<sup>29</sup>.

Beyond providing technical validation data or examples, a key component of generating buy-in is visualization. Visualization is not a mere aesthetic choice of the DSM designers, but rather an important aspect that impacts not only how DSM will be received and used, but also what decisions are reached. Previous research has showed that changing how data or models are displayed can significantly affect risk aversion/acceptance<sup>40</sup>, ability to come to a negotiated agreement between stakeholders<sup>41</sup>, willingness to use a tool<sup>42</sup>, and ability to apply presented knowledge<sup>43</sup>. As a result, attention must be paid to how the DSM is presented. This does not necessarily mean that a universal DSM visualization standard should be developed (though it could), as industry partners may be better served by using different visualization schemes in-house than those used by DoD. If there is no standard visualization method, however, differences must be made clear, as these could affect decision-making and lead to conflicts in assessments of model results. It should be noted that the importance of visualization applies not only to DSM itself, but also to the reference materials made available to the users for training, as previous research has shown that even the layout of brochures can have significant impact on information gained<sup>44</sup>. As a result, multiple manuals covering the same model may be necessary if different types of users will be interacting with the model.

**Model Comparisons/Validity.** If a heterogeneous model structure is chosen, as is currently preferred by the DoD, an issue of model comparisons arises. If during the bidding competition, two firms provide DSMs of their preliminary design, but are using different modeling software within each DSM, there may be some difficulty in comparing the simulation results of the two DSMs. This is not necessarily a new issue to the DoD, as claims of competing firms already need to be evaluated in the current acquisition procedure, but it should be acknowledged that this problem will not vanish under the new digital framework.

---

<sup>40</sup> S. Park, L. Rothrock. Systematic analysis of framing bias in missile defense: Implications toward visualization design. *European Journal of Operational Research* 2007; vol. 182, 1383-1398.

<sup>41</sup> M. E. Fitzgerald, A. M. Ross. Effects of Enhanced Multi-party Tradespace Visualization on a Two-person Negotiation. *Procedia Computer Science* 2015; vol. 44, 466-475.

<sup>42</sup> K. Wodzicki, E. Schwämmlein, U. Cress, and J. Kimmerle. Does the Type of Anonymity Matter? The Impact of Visualization on Information Sharing in Online Groups. *Cyberpsychology, Behavior, and Social Networking* 2010; vol. 14, 157-160.

<sup>43</sup> A. L. Alexander, T. Brunyé, J. Sidman, and S. A. Weil. From gaming to training: A review of studies on fidelity, immersion, presence, and buy-in and their effects on transfer in pc-based simulations and games. *DARWARS Training Impact Group* 2005; vol. 5, 1-14

<sup>44</sup> W. G. Morgan Baruch Fischhoff, A. Bostrom, L. Lave, and C. Atman. ES&T Features: Communicating Risk to the Public - First, Learn what people know and believe. *Environmental Science & Technology* 1992; vol. 26, 2048-2056.

One potential method of diffusing IP disputes is for each individual model in DSM to be treated as a black box by all participants except for the owner of the component and the DoD. This however introduces evaluation issues that may hamper the design process. Specifically, it falls upon the DoD to ensure compatibility of calculation methods between the models as the firms will be unable to analyze each other's models themselves. This will result in bugs being harder to identify and could result in erroneous model results, jeopardizing both the technical validity of the DSM and the trust that the firms have in it as well, reducing buy-in.

### **Potential DSM Enabling Structures**

With these issues in mind, it is worthwhile to consider the enabling structure of DSM along two different dimensions. The first is how the modeling software packages are developed and the second is how the modeling software packages are used. As stated earlier, these options are not intended to be an exhaustive list, but should provide grounding for consideration of IP and KA issues in the design of DSM.

#### ***Model Package Development.***

In this first dimension, the primary relevant issues are the IP of the modeling software packages (but not the products being modeled) and KA of model comparisons. Three general options exist. The first, and most simple option, is that the DoD develops its own modeling packages for each part of the DSM and then mandates that industry partners use these packages. This could be an extension, though a highly ambitious one, of the CREATE program<sup>24</sup>. This homogenous structure would eliminate any IP disputes over the modeling packages themselves and eliminate any KA issues of resolving inter-model differences. Furthermore, this would alleviate fears of losing access to DSMs of older systems since the completed DSM would not be reliant on any privately-held software. Doing this, however, would fail to take advantage of the large amount of intellectual capital present in the private defense industry and would require the DoD to continually maintain and improve these software packages.

The second option would be for the DoD to standardize the model interface akin to the typical standards development method discussed earlier. This heterogeneous structure would allow each firm to create its own modeling packages (or license those of others as it chose). This option does not directly resolve any IP-sharing issue (that depends on the model usage dimension discussed in section 0) nor does it resolve the model validity issues. Unless the software packages are permanently licensed or outright purchased, this option raises the possibility of the DoD not having perpetual access to all components of the DSM, especially if a firm goes out of business or fails to maintain its own modeling packages. It should be mentioned that this is currently the structure favored by the DoD, due to its combination of modularity and openness<sup>45</sup>.

The third option is a hybrid of the first two, being privately-developed like the second option, but homogenous like the first. Here the industry partners would be relied upon to generate the modeling packages and these packages would be required to interface with one another, but the

---

<sup>45</sup> P. Zimmerman. Modularity and Open Systems: Meaningful Distinctions. 18th Annual NDIA Systems Engineering Conference; 2015

DoD would select specific modeling packages to be used for each portion of the DSM, likely in some sort of periodic bidding process (an acquisition project in its own right). The DoD would not just buy or license the modeling software for itself, but also for each other industry partner, otherwise the owner of the selected software package would have effective monopoly power over other members of the industry. This option does not address perpetual access issue any more than the second option, but does resolve the inter-model comparison issue while still relying on the industry to generate and improve modeling packages.

These options, along with their pros and cons, are summarized in Table 1.

Table 1. Summary of Model Package Development Structures

	<b>Pros</b>	<b>Cons</b>
<b>DoD Developed</b>	Reduces IP disputes Can maintain access Eliminates inter-model comparisons	Does not utilize industry expertise Requires DoD to maintain and update
<b>Heterogeneous, Privately-Developed</b>	Fully leverages competitive industry Minimizes DoD effort	Does not resolve IP disputes directly Requires inter-model comparison Potential lack of continued access
<b>Homogenous, Privately-Developed</b>	Reduces IP disputes Partially leverages competitive industry Reduces DoD effort Eliminates inter-model comparisons	Introduces miniature monopolies Potential lack of continued access Does not fully leverage industry expertise Does not minimize DoD effort

**Model Use.** The second dimension, once the individual modeling packages are selected, concerns how they work together in practice to form the DSM. Once again, three general options exist, two integrated and one distributed, though there different variations of each.

The first option is for the complete DSM is held by one party. This could be DoD for the entire system lifecycle or held by the lead industry partner during the design process before transferring to the DoD after design. This method could potentially reduce cybersecurity concerns as the DSM would not necessarily need to be connected to any network or the internet. If the DSM is held by the lead industry partner, however, it fails to address the IP-sharing issues. Regardless of who, specifically, holds the complete DSM, this option may reduce the benefits of the DSM. The design process may be inhibited, as it may be difficult and time-consuming for one central authority to keep the DSM updated, continually run simulations, and distribute the results to the other industry partners.

The second option is to have multiple copies of the complete DSM, each held by the different firms and organizations involved in the product development process. This option addresses some of the design inhibition concerns of the first option by allowing various firms to run their own simulations. This comes at the cost of introducing additional cybersecurity concerns in whatever syncing process is used between the various copies. Furthermore, as has been discussed by other researchers<sup>32</sup>, the DSM is going to be a massive model, in terms of computational power necessary to run simulations with it. It is fully possible that only a select few industry partners would have the computing capability to run the full DSM themselves. In the case of particularly complicated systems, it is possible that only the DoD itself would have such capability. This may limit the benefits of this option in comparison to the first option, though the industry partners would still likely be able to run the DSM either in part or at lower resolutions, while relying on the DoD for the higher resolution, full system simulations.

The third option is for each firm to maintain control over the models they create and have the DSM exist as a sort of distributed network, similar to the FDA's Sentinel Initiative. Cybersecurity is once again of high concern with such an arrangement and the distributed system may slow down simulation run times, particularly if multiple parties simulate the system at the same time. This option could provide for more robust protections of IP of the industry partners. It is likely that if this structure was chosen, it would only exist prior to hand-off to the DoD, as the DoD would like to maintain the ability to run its own simulations without relying on an entire network of other parties.

Each of these options, summarized in the table, could be implemented with various degrees of IP protections of the modeling software and the models themselves. In either distributed or integrated form, models and their associated software could be handed off as black boxes working in the kind of structure described in the previous section. This could help alleviate concerns over sharing IP with other industry members but would come at the cost of placing the burden of assessing inter-model compatibility and model comparisons on the DoD, with who, presumably, full details on the software and models would be shared so these assessments could be made.

Table 2. Summary of Model Use Structures

	<b>Pros</b>	<b>Cons</b>
<b>Centralized – Single Copy</b>	Reduces IP disputes	Difficulty in updating
	Reduces security risk	Hampers iterative design
<b>Centralized – Multiple Copies</b>	Allows for iterative design	Few firms can host full DSM Increases security risk Does not address IP disputes
<b>Distributed</b>	Reduces IP disputes	Increases Cybersecurity issues
	Allows for iterative design	Requires transition to centralized during hand-off to DoD Potentially technically difficult Increases simulation-run times Increases security risk

## Discussion

In the near term, some concrete and more immediate steps concerning non-technical issues could be pursued to advance DSM. For instance, there has been some confusion in understanding DSM due to the lack of clear distinction between DSM and DTh in previous literature<sup>32</sup>. It may be worthwhile for future official explanations of the concepts to be more specific about their differences, and perhaps include examples of what is a part of each. Alternately, it may be possible to combine the terms if the concepts referred to are too intertwined to separate.

Investigation of the nontechnical issues surrounding DSM raises some wide-reaching issues requiring further research. For instance, increased awareness of non-technical challenges supports the case for further investigation by legal experts concerning the potential exception of DSM development from antitrust laws. If the IP used to develop DSM does not need to be licensed under FRAND terms, then additional options are open to the DoD for the enabling structure of the DSM. Regardless of the legality of such options, pursuing an approach that is not under FRAND terms may result in political and/or industry pushback.

In the DoD context, the government serves several roles: customer (thereby supplying financial incentive), the standards enforcer (regulatory incentive), and neutral mediator (serving as non-competing party with whom information can be shared). These and other roles and responsibilities for handling the non-technical aspects of the digital framework require further investigation. Generating use cases related to all facets of creating, sustaining and protecting the digital models is an open area of investigation.

Evolving enabling structures for successful implementation of the digital framework requires investigation in several directions. Looking further into effective model package development structures and model use structures that have been successful in other domains may provide insights. There is a need for more research on visualization schema and standardization, including how these are applied throughout development of the digital framework. This may require

comparative testing of multiple display styles and novel research if existing methods of visualization are insufficient for the systems being designed and managed by DoD. Additionally, ongoing research concerning the technical challenges needs to be effectively linked to investigation into the entangled non-technical challenges.

Many assessments of DSM's potential benefits to acquisition and operations assume that the IP and KA issues discussed would be perfectly resolved, though these same assessments often did not provide means of achieving such resolutions. While these issues are not insurmountable, they require no less attention than the various technical challenges facing DSM. These issues will not resolve themselves; issues such as user buy-in could scuttle the DSM entirely, even if technical challenges are overcome. This is not to be pessimistic, however, as there are some viable paths forward for investigating these issues and finding strategies for achieving the goal of next generation model-centric engineering. The findings of this study suggest the need for further awareness of non-technical challenges, and further investigation into practical strategies and enabling structures for realizing the promise of digital models.

### **Next Steps**

The investigation of non-technical challenges for digital system models has revealed a number of questions that require further definition and research. The efforts of this study have been used to inform research activities in the IMCSE research roadmap. Additional investigation will be performed in the next phase of research on curation of model-centric environments.

---

## MODEL CURATION

The 2015 IMCSE Pathfinder Workshop identified research needs from both a model-centric perspective and an interactive perspective<sup>46 47 48</sup>. Model curation was cited as an important topic for investigation in evolving model-centric engineering that involves both these perspectives.

### Importance of Model Curation

Given a shift in engineering to a model-centric paradigm, there is a need to better understand *where* the role of the human (versus automation or “AI”) is essential in the effective management and utilization of model environments. This includes the models (managed as organizational assets), supporting infrastructure, and the associated protocols and practices. Digitized legacy systems and new digital system models will provide the basis for designing and evolving systems into the future. This drives the criticality of models as assets and necessitates change in model-related policy and practices. Accordingly, there is an urgent need to mature a practice of “model curation” including a “model curator” functional role within engineering organization.

Model curation has been raised as a topic of growing importance within the systems community. Rouse<sup>49</sup> stresses that the wealth of existing models is often not used because of a lack of knowledge of these resources and the difficulty in accessing them. From a practical perspective he says, “Professionals in modeling and simulation seldom access academic journals reporting the models they use. Even when practitioners are aware of such publications, they are seeking computer codes, not research treatises”. To bridge this gap, there needs to be a single point of access to this body of knowledge, enabled with downloading computer codes, documentation on assumptions and use, and dissertations on the development and validation of these models.

The participants in the workshop agreed that there is no consistent meaning across or between the “sea” of models that are used or exist. There is a lack of precise semantics, especially in the behavior and timing/interactions of models. Many organizations are beginning to develop repositories of models and archive associated data, which is one of the fundamental activities in model curation. Repositories support re-use of building blocks from previous analyses and tailoring to the specific requirements of one’s problem. Potential benefits include improving the quality and timeliness of analysis, by developing a repeatable framework providing guidelines for accessing, designing, and implementing models.

Evolving the science of model curation will involve a partnership of government, industry and academia. There are many challenges and interesting problems related to this topic, and there is great potential benefit for practitioners, researchers and educators. As shown in the table below, each contributes in support of a model curation function and each can benefit in a number of

---

<sup>46</sup> D. H. Rhodes and A. M. Ross, Interactive Model-Centric Systems Engineering Technical Report SERC-2014-TR 048-2, February 28, 2015.

<sup>47</sup> M. Blackburn, R. Cloutier, E. Hole, M. Bone, Transforming System Engineering through Model-Centric Engineering, Technical Report SERC-2015-TR-044-3, 2015.

<sup>48</sup> A. M. Ross, Interactive Model-Centric Systems Engineering: Progress and Updates, Presentation at 6th Annual SERC Sponsor Research Review, 2014.

<sup>49</sup> W. B. Rouse, Modeling and Visualization of Complex Systems and Enterprises, Exploration of Physical, Human, Economic and Social Phenomena, 2015.

ways. Further consideration is needed to understand the role and impact of a model curator within these communities.

	Example Benefits	Example Contributions
Practitioners	<ul style="list-style-type: none"> <li>• Have access to state-of-the art research in modeling and simulation</li> <li>• Have timely access to validated modeling resources and support from curation specialists</li> <li>• Evaluate different modeling alternatives before investing heavily in one</li> </ul>	<ul style="list-style-type: none"> <li>• Submit validated models and data to in-house and/or shared repositories</li> <li>• Propose model improvement based on feedback from real use cases implementations</li> <li>• Work towards community-shared semantics in model-centric activities</li> </ul>
Researchers and Educators	<ul style="list-style-type: none"> <li>• Have research opportunities to work on interesting model-centric topics</li> <li>• Structured and quality reviewed resource for students (e.g., apply modeling for class projects)</li> </ul>	<ul style="list-style-type: none"> <li>• State-of-the-art research and impact on modeling and simulation tools and methods</li> <li>• Strengthen the theoretical grounds of model curation and human-model interaction</li> </ul>

As engineering practice becomes increasingly model-centric, models are valuable assets for designing and evolving systems. The need for model curation accordingly becomes a necessary functional role in organizations. This is a relatively new idea in the systems community, though progress has been made on some of the activities involved in curation. Maturing an approach for model curation in the systems engineering field can leverage the work of other related curation practices.

Digital curation is closely related to modeling curation, there is much to borrowed and adapted from this practice. Practices on collaboratively developed model repositories and their management provide additional insights for model curation<sup>50</sup>. Another related area is social content curation, focusing on collaborative sharing of Web content organized around one or more particular themes or topics. “Social curation can be defined as the human process of remixing social media content for the purpose of further consumption”<sup>51</sup>. It is viewed as a complement to traditional data exploitation methods such as algorithmic search and aggregation.

Curation of institutional collections, such as the role of the traditional museum curator, offers an analogy for the many potential responsibilities of the model curator. As with a large institution, the curation function is carried out by a team of individuals and the same would be true of the model curation in the systems field. The museum curator’s role is an essential one where curators oversee collections of artwork and historic items, with support from archivists who appraise, edit,

<sup>50</sup> European Bioinformatics Institute (EMBL-EBI), <http://www.ebi.ac.uk>, accessed 21 October 2015

<sup>51</sup> Duh, K., Hirao, T., Kimura, A., Ishiguro, K., Iwata, T., & Yeung, C. M. A. (2012, May). Creating Stories: Social Curation of Twitter Messages. In ICWSM.

and maintain permanent records and historically valuable documents. Specialists and technicians are also involved in various capacities.

Extending from the various types of curation roles and activities of other fields, the model curator's role is envisioned to include a number of major responsibilities and support of various staff. The model curator (curation function) would set and administer model-related policies and practices. The curator would ensure models and related documents are authenticated, preserved, classified and organized accordingly with model metadata standards, to be defined. The curator may own the data management for models and related information, or oversee the ownership by other individuals or organizations. As needed, a curator would meet with individuals and teams, who will create, use and re-use models, helping to determine a useful classification of individual and sets of models. At the organization level, the curator may organize training and special projects related to model-based engineering. The curator may also participate in creating and maintaining model-based work environments.

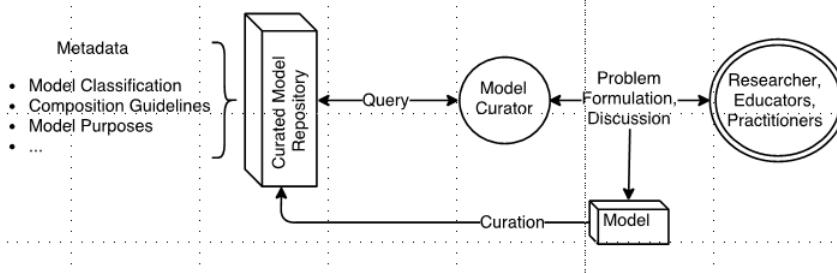
The model curator role needs to become a formal role in engineering enterprises given the needs and challenges in model-centric engineering. Curators will need to know about a number of things including model ontologies, model meta-data, latest modeling techniques and classes of models, policies on data rights, code of ethics, and others. Effective model curation necessitates clarity across the systems community in characterizing and handling models. It requires formalizing knowledge of models and determining a distinctive set of model characteristics (purpose, input/output types, logic, assumption types, model incompatibilities, etc.). A model classification drawn from prior use is a first step towards generating a framework for curating models. There are many facets of model curation needing further investigation and elaboration.

Successful model curation will require a shared understanding of numerous aspects of models (Figure 4) within the systems community, such as model purpose, model classification, model characteristics, and model composition guidelines.

## Managing models as assets, to improve model-centric systems engineering practices and research

- Research needed on modelling metadata: model ontologies, modelling techniques, data rights, code of ethics, work environments, characteristics of models, composition of hybrid modelling techniques, and more
- Curator function: to authenticate, preserve, classify models; to be the keeper and advocate for the curated database of models
- Effective and timely matching of client problem needs and constraints with adequate and insightful individual, or sets of, models

*A model curation function would require contributions from, and provide benefits, to the wider modeling community: practitioners, researchers, and educators ... as done in digital curation and social curation practices.*



Reymondet, L., Rhodes, D.H. and Ross, A.M., Considerations for Model Curation in Model-Centric Systems Engineering, IEEE SysCon 2016

Figure 4. Importance of a Model Curation Function

### Model Purpose

The selection of an appropriate modeling approach is determined by the question or problem being addressed. It is recognized that a model is not a full representation of the system. A model should be a useful representation, and draws its goodness from being simple, yet evocative<sup>52 53</sup>. Therefore, the purpose for modeling sets the boundary for model depth and breadth.

Various types of models have been enumerated by the systems engineering community, but there appears to be insufficient attention given to model purpose itself. In the field of informatics, McBurney<sup>54</sup> proposed nine model purposes, from understanding, predicting or controlling natural reality (e.g. Newton's laws) to playing and enabling the exercise of human intelligence, ingenuity and creativity, in developing and exploring the model (e.g. so-called serious games).

When tackling the complexity of a technical system such as an aircraft, modelers have the intent to classify, explain, predict, control, detect or diagnose the system's input/output behavior. As for human behavioral and social phenomena, different modeling approaches are appropriate at the people, process, organization and ecosystem scales<sup>4</sup>. For example, at both ends of the

<sup>52</sup> G. E. P. Box, Science and Statistics, Journal of the American Statistical Association, 1976.

<sup>53</sup> S. Banks, Exploratory Modeling for Policy Analysis, in Operations Research, Vol. 41, No. 3, 1993.

<sup>54</sup> P. McBurney, What are Models for?, in: M. Cossentino, K. Tuyls and G. Weiss (Editors): Post-Proceedings of the Ninth European Workshop on Multi-Agent Systems (EUMAS 2011). Lecture Notes in Computer Science, volume 7541. Berlin, Germany: Springer, pages 175- 188

spectrum, models at the people level will seek to detect, explain and resolve anomalies in human behaviors (e.g., consumer behavior), while models at the ecosystem level will seek to detect, explain and resolve anomalies in societal behaviors (e.g. intra-firm competitive relations). Lower-level model outputs serve as inputs for higher-level models and higher-level outputs are constraints on lower levels of modeling.

Zacharias et al.<sup>55</sup> point out that individual, organizational and societal models, do not predict exactly what humans will do, as individuals or in groups, but rather help forecast a range of potential action outcomes, draw attention to potential unintended consequences, and highlight variables that are overlooked in a particular situation. Accordingly, model purposes include: to analyze fragmented information and develop courses of action based on the likelihood of desired outcomes; to train personnel, simulating the environment, dynamics and providing performance feedback; and to design and evaluate a technical system, predict its performance and make decisions based on cost-benefit tradeoffs.

A shared understanding of model purposes would facilitate the critical transition from the motivation for modeling (e.g. perceived complexity in a system, or issues perceived in an organization), to the choice of an adequate set of models (e.g. to tame the system complexity, or to help decide upon management strategies). Empirical observations suggest the manner in which this is performed today is mostly ad hoc, or biased by practice anchoring within a field or a group.

### **Model Classification**

The systems community has instantiated many different models to support designing, evaluating, and testing technical systems (requirements/functional diagram, computer-aided design, design structure matrix, multi-attribute trade space exploration, cross impact analysis, fault tree analysis, hardware-in-the-loop/human-in-the-loop simulation...), analyzing social networks and phenomena (system dynamics, queuing models, agent based models, graph theory models...), and understanding human behavior (mental models, Petri net models, task models, microeconomics models...) among others.

A suggested first step in making the transition from the motivation for, to the selection of, a model would be to examine one's issue and expectations from the model against previous instances. In this endeavor, a modeler would use a classification of previously adopted modeling approaches along several characteristics, including model purpose. It is also expected that such a survey would highlight limitations of each modeling approach in its context of use, which are equally important as the purpose.

Model characteristics would provide additional elaboration to the metadata that is needed in organizing the envisioned classification, to enable case-specific selection of models. Some example characteristics could be input/output type, software environment and computational

---

<sup>55</sup> G. L. Zacharias, J. MacMillan, S. B. Van Hemel, Behavioral Modeling and Simulation, 2008.

language used, social and technical modeling scales (from agents to societies, or components to systems), and treatment of time (continuous/discrete). Establishing a useful list of characteristics will require contributions from practitioner and research communities.

### **Model Selection**

Traditional, often computational, methods for systems engineering are inadequate for a complex system such as a sociotechnical system, which includes soft and hard problems, a very large number of technical components and autonomous actors, and emergent behavior created by soft-hard, social-technical, non-linear interactions. Using a set of models enables insight into various aspects of system complexity.

Bankes<sup>9</sup> opposes consolidative computational modeling of deterministic systems (models as surrogates for real systems) to exploratory modeling of systems plagued with uncertainty and unknown unknowns (models as means of testing hypotheses and exploring ranges of possible outcomes). It is argued that exploratory modeling can only produce useful results through a constellation of alternative models. By using multiple simple models, complexity is exported outside the models to the ensemble of model outcomes, from which modelers and stakeholders must make sense.

Selecting one or several models for a complex case study requires breaking down the problem into scoped questions such that each question translates into a model purpose. For example, designing an urban smart power grid is a complex problem taken as a whole. It could be broken down into the following scoped problems (1) Tradeoff energy production mix design alternatives, predict reliability and vulnerability of each design against dynamic demand loads (technical level); (2) Predict consumer preferences regarding energy bills, effort involved, and value attached to contributing to environmental sustainability (people level); and (3) Optimize dynamic pricing of utilities – organization level - within the rules set by local government – environment. Rouse<sup>4</sup> proposes such a multi-level modeling approach for dealing with complex socio technical systems. Each problem would then call for a tailored modeling approach, for example: 1a.Trade space exploration, 1b.Fault tree analysis, 2.Multi-attribute utility functions, and 3.Microeconomics model of key stakeholders. This model set (1a, 1b, 2 and 3) is neither unique, nor exhaustive and a modeler might only be interested in one aspect of the problem. Therefore, identifying the problem clearly and formulating a clear model purpose is essential in setting boundaries for the modeling effort, and could be assisted by a model curating function.

## **Model Composition Guidelines**

Recent work on hybrid modeling<sup>56 57 58 59</sup> and multi-scale modeling<sup>4</sup> point towards the usefulness of using not a single but a set of models to study a complex system. LaTour<sup>12</sup> implements a system dynamics model interactively with a trade space exploration model for investigating the impact of time on the lifecycle and procurement of GPS satellites. Mathieu et al.<sup>13 14</sup> implement a system dynamics model, a Petri net process model and an agent-based model for simulating the response timeliness and effectiveness of the Air and Space Operations Center to a series of critical events, at the mission, process and operator levels. Zulkepli<sup>15</sup> implements a system dynamics model and a discrete-event simulation to simulate the interactions between healthcare personnel stress, and patient non-recovery and readmission rates. Synergies between models are seen to increase modeling capabilities and insights into problem solving while reducing the limitations of individual techniques.

Solving the original overarching problem requires being able to integrate different models so as to make sense of the whole. Issues arise especially with computational models, when the integration is automated. Assumption consistency between models and between models and the reality they represent is one issue<sup>4</sup>. Additionally, interfacing two models, especially two computational models, raises issues such as incompatibilities in data types, in naming schemes, in logic mechanisms. Model composition may consist in parallel models - run independently from one another – or hybrid models - coupled such that the output of one serves as the input for another, and vice versa – and the development, testing and validation methods differ between the two compositions. However, the risk in aggregating many models is to end up with a model of high resolution but low practical utility and transparency.<sup>9</sup> Hence, guidance for model composition is needed in model curation, in particular for modeling complex systems, where multi-level modeling approaches are useful.

## **Discussion**

Further research is needed as well as exchanges between academic and practitioner communities on the form that model curation should take in systems engineering. The above mentioned axes of research – model characterization, purposes, classification, selection and composition guidelines – are essential to build a shared understanding of what a model is before attempting at creating a large-scale curated model repository. In-depth investigations of previous model use cases enabled potentially enabled by social model curation could help consolidate a repository of modeling approaches from the bottom-up, while academic research can provide a theoretical basis and top-down structure for a database of models.

---

<sup>56</sup> P.A. La Tour, D. E. Hastings, Combining System Dynamics with Tradespace Exploration to Explore Future Space Architectures, 2015.

<sup>57</sup> J. Mathieu, J. Melhuish, J. James, P. Mahoney, L. Boiney, B. White, Multi-Scale Modeling of the Air and Space Operations Center, The MITRE Corporation, 2007.

<sup>58</sup> J. Mathieu, J. James, P. Mahoney, L. Boiney, R. Hubbard, B. White, Hybrid Systems Dynamic, Petri Net, and Agent-Based Modeling of the Air and Space Operations Center, The MITRE Corporation, 2007

<sup>59</sup> J. Zulkepli, T. Eldabi, N. Mustafee, Hybrid simulation for modelling large systems: an example of integrated care model, Proceedings of the 2012 Winter Simulation Conference, 2012.

## **Next Steps**

Further dialogue within the system engineering community is needed to determine the usefulness and limitations of a model curator function in model-centric systems engineering. In the next phase of IMCSE research, there will be a more intensive study on curation for model-centric environments, engaging experts from industry, government and academia. The goal will be to develop a comprehensive roles and responsibilities description, and gather findings that may support the development of a model-centric environment self-assessment.

---

## **COGNITIVE AND PERCEPTUAL CONSIDERATIONS IN HUMAN-MODEL INTERACTION**

Over the past several decades the work environment of the engineer has shifted from a hands-on workbench type of environment to model-centric work stations and collaborative laboratory environments (Figure 5). As model-based engineering continues to evolve, engineers will increasingly work with many types of models, ranging from highly abstracted representations to realistic multi-dimensional model. And, systems of the future may have digital twins, a model-centric replica of the operational system. Advances in the technology and computational resources have been steadily made, and the laboratory environments have become increasingly sophisticated. Yet, the many facets of the human-model interaction experience remain relatively unexplored. Learning from past situations with similar considerations is a useful place to start in investing the human aspects.



**Figure 5. Shift from the ‘workbench’ to model-centric environment**

---

## **BACKGROUND**

An invited workshop held in January 2015 seeded a research agenda around the topic of human-model interaction, identifying research needs from both a model-centric perspective and an interactive perspective. Participants agreed that progress has been made on standards, methods and techniques for model-based systems engineering, yet little attention has been given to human-model interaction (Rhodes and Ross 2015)<sup>17</sup>Error! Bookmark not defined..

Related fields addressing humans and systems/technologies have existed for some time. A science of human-systems integration (HSI) has emerged, but focuses on operational systems. Human-computer interaction (HCI) is another similar field with useful knowledge and principles extensible to IMCSE, but the focus tends toward the design of displays. IMCSE is primarily concerned with the interaction with models, as abstractions, and the role of models in making decisions.

---

## **GLASS COCKPIT INVESTIGATION<sup>60</sup>**

With the assumption that learning from past situations with similar considerations is a useful place to start investigating aspects found in human-model interaction. The research team investigated the case of the shift from traditional aircraft cockpits to glass cockpits. Within aircraft, cockpit displays present pilots with models of the aircraft’s state in order to facilitate

---

<sup>60</sup> German, E.S. and Rhodes, D.H., “Human-Model Interactivity: What can be learned from the experience of pilots transitioning to glass cockpit?”, Conference on Systems Engineering Research, March 2016.

appropriate decision making and action. The goal of exploring transition of modern aircraft from use of analog to digital “glass cockpit” displays was to draw out lessons learned from the transition’s impact on human and system performance. Through investigation of case studies of aircraft accidents and subsequent research findings, areas of concern in the interaction between glass cockpits and human pilots have been identified. Significant research has been performed on aircraft accidents to not only retroactively address accidents, but also to identify areas susceptible to failure and to determine the causes of these failures, with an end goal of mitigating future occurrences of accidents. Operating on the premise that the cognitive and perceptual issues found in the cockpit transcend to broader terms of human-model interaction, the research team investigated these lessons to enhance initial thinking into the role of human-model interaction within the emerging field of interactive model-centric engineering.

The term “glass cockpits” began making its way into the aviation community in the 1970s with the transition from electromechanical instruments to electronic flight displays. Used initially to describe displays incorporating cathode ray tubes, “glass cockpit” has since evolved into a descriptor for digital flight displays and automation systems within aircraft in general (NTSB 2010)<sup>61</sup>. The arrival of the glass cockpit equipped Boeing 757 and 767 in the early 1980s ushered in the use of glass cockpit and automation technology within commercial aviation, progressing to become standard design in nearly all modern aircraft (Stauch 1998)<sup>62</sup>. This new technology sought to improve system functionality by increasing human capability and efficiency through automation of flight operations and ultimately allowed the crew composition of commercial aircraft to be reduced from three to two members (Weiner 1989)<sup>63</sup>. As noted by Endsley (1996)<sup>64</sup>, however, these benefits from automation also came with changing the pilot’s role from flying to monitoring an automated system, a “role people are not ideally suited to”. Analysis of aircraft accident case studies provides insight into challenges that glass cockpits and their automation have caused for pilots within the cockpit environment.



<sup>61</sup> NTSB, *Introduction of Glass Cockpit Avionics into Light Aircraft*. National Transportation Safety Board. March 9, 2010.

<sup>62</sup> Strauch, Barry. “Projected Flight Path Displays and Controlled Flight Into Terrain Accidents.” *Digital Avionics Systems Conference*. Bellevue, WA: IEEE, 1998. E43/1 - E43/8.

<sup>63</sup> Wiener, Earl L. *Human Factors of Advanced Technology (“Glass Cockpit”) Transport Aircraft*. Moffett Field: National Aeronautics and Space Administration, 1989.

<sup>64</sup> Endsley, Mica R. “Automation and Situation Awareness.” In *Automation and Human Performance: Theory and Applications*, 163-181. Mahwah: Lawrence Erlbaum, 1996.

**Figure 6. Displays used by pilot (a) traditional, analog cockpit; (b) glass cockpit.**

### **Case-based Investigation**

The research team investigated three historical cases to examine cognitive and perceptual considerations in situations where humans are interacting with high technology environments that include abstract and synthesized information display.

**Nagoya.** A first case example concerns an airline crash in April 26<sup>th</sup>, 1994. While piloting an Airbus A300-600 on landing approach in Nagoya, Japan the First Officer (FO) mistakenly engaged the Go-Around mode as the aircraft neared 1000ft above ground level. The aircraft appropriately responded by autonomously adding power and initiating a climb that the FO tried to manually counteract to keep the plane on the appropriate glide path. While the Captain noticed the erroneous initiation of the Go-Around mode and told the FO to disengage the mode, the FO failed to do so. The FO managed to stop the plane's ascent and the autopilot was engaged, but 19 seconds later the autopilot caused the plane to pitch up again, which caused the FO to disengage the autopilot. Around 570ft the aircraft sensed near-stall conditions, autonomously staged a stall recovery and began climbing once again, yet this time the pilots were unable to stop the climb which ultimately led to stall, inadequate time for recovery, and a tail-first crash landing killing 264 of the 271 individuals onboard (Baxter et al. 2010)<sup>65</sup>.

**Strasbourg.** The next case example occurred in Strasbourg, France on January 20, 1992 when an Airbus 320 impacted the ground while on descent for landing. Prior to beginning the approach, the crew received last minute instructions from Air Traffic Control to complete a straight in landing rather than the expected circling approach. This unanticipated guidance resulted in an increased workload for the crew as they worked to complete the preparations for landing in an earlier than expected manner, yet they continued with the directive to execute a straight in approach. As part of the preparations, the crew entered the number "33" into the flight computer to set the appropriate glide path angle of -3.3 degrees. They failed to realize, however, that the computer's mode was set to rate of descent and that they actually commanded the aircraft to descend at 3,300 ft/min. The aircraft proceeded as was mistakenly directed and subsequently crashed into the ground, with only nine of the 96 individuals onboard surviving (Moshier et al. 2007)<sup>66</sup>.

**Cali.** Reminiscent to the event at Strasbourg, an American Airlines Boeing B757 received Air Traffic Control guidance on December 20, 1995 to complete an unplanned, straight-in landing approach for its destination, Cali, Columbia. Needing to adjust their flight plan to complete the approach, the crew proceeded to enter in the next appropriate navigation waypoint, "ROZO," into the flight computer. However, after inputting "RO" the waypoint "ROMEO" was the first available point on the list which the crew mistakenly selected; the aircraft then began navigating to a waypoint located 132 miles away from the destination. Approximately a minute following

---

<sup>65</sup> Baxter, Gordon, Denis Besnard, and Dominic Riley. "Cognitive mismatches in the cockpit: Will they ever be a thing of the past?" *Applied Ergonomics*, 2010: 417-423.

<sup>66</sup> Mosier, Kathleen L, Nikita Sethi, Shane McCauley, Len Khoo, and Judith M Orasanu. "What You Don't Know Can Hurt You: Factors Impacting Diagnosis in the Automated Cockpit." *Human Factors*, 2007: 300-317.

the plane's course adjustment away from Cali, the crew realized their mistake and reprogrammed the flight to the appropriate point, ROZO. Assuming the situation rectified, the crew failed to realize the deviation from the original flight path set the airliner on a collision course with a mountainside. Only 4 individuals out of 163 survived the crash (Basard and Baxter 2006)<sup>67</sup>.

The historical cases underline a theme of similarity (Mosier et al. 2009)<sup>68</sup> in that they would not have occurred in the absence of highly automated equipment within the cockpit, demonstrating a breakdown of human interactivity with the aircraft that ultimately led to devastating results. This evidence suggests further investigation of the causes and potential mitigations of these types of errors presents an opportunity to gain insight into human-model interactivity, specifically in highly automated environments like those found in aircraft.

### Cognitive Coherence

Transition of aircraft cockpit technology has largely changed the role of the pilot from one that requires "stick-and-rudder" skills, to one primarily concerned with programming and monitoring the aircraft's automation (Mosier et al. 2009)<sup>68</sup>. As described by Mosier, this shift in the pilot's role has also accentuated the importance of coherence competence: "an individual's ability to maintain logical consistency in diagnoses, judgments, or decisions"(Mosier et al. 2001)<sup>69</sup>. The displays within aircraft present nearly all of the necessary data to safely fly the plane, and if the pilot can maintain coherence and take appropriate action throughout the entirety of the flight then the pilot has succeeded. Mosier also notes that many piloting errors manifest themselves as failures of coherence in that they fail "to note or analyze important information in the electronic 'story' that is not consistent with the rest of the picture"<sup>66</sup>. The outcomes these cases all resulted from a failure to maintain coherence throughout the entire flight. While maintaining coherence is a primary objective for pilots, there are many means through which automation can contribute to the breakdown of effective coherence.

**Automation Bias.** Mosier and Skitka (1999)<sup>71</sup> define automation bias as "the use of automation as a heuristic replacement for vigilant information seeking and processing," which can result in commission errors (incorrectly following an unverified automation directive) and omission errors (failing to identify an issue not identified by an autonomous system). An everyday example of a commission error would be a driver blindly following a GPS navigation's incorrect directive to turn the wrong way onto a one way street. Additionally, missing the proper highway exit due to lack of warning from the navigation system would constitute an error of omission (Parasuraman and Manzey 2010)<sup>72</sup>. Specifically related to human interaction with automated decision aids, automation bias seems to be influenced by three different factors. First, humans often choose to proceed down the path of least cognitive effort. Humans also exhibit a tendency to perceive

<sup>67</sup> Besnard, Denis, and Gordon Baxter. "Cognitive Conflicts in Dynamic Systems." In *Structure for Dependability: Computer-Based Systems from an Interdisciplinary Perspective*, 107-124. London: Springer London, 2006.

<sup>68</sup> Mosier, Kathleen L, Linda J Skitka, Susan Heers, and Mark D Burdick. "Automation Bias: Decision Making and Performance in High-Tech Cockpits." *The International Journal of Aviation Psychology*, 2009: 47-63

<sup>69</sup> Mosier, Kathleen L, Linda J Skitka, Melisa Dunbar, and Lori McDonnell. "Aircrews and Automation Bias: The Advantages of Teamwork?" *The International Journal of Aviation Psychology* 11, no. 1 (2001): 1-14

automated decision making and performance as superior to their own, leading to an overestimated trust that the system is performing appropriately for the given situation. A third factor influencing automation bias is the phenomena of perceiving automated aids as fellow crew members and diffusing responsibility. This can lead to a “social loafing” behavior where human operators perceive themselves as less responsible for the system performance and outcome (Parasuraman and Manzey 2010)<sup>72</sup>. The accidents at Strasbourg and Cali offer examples that manifest potential instances of automation bias. At Strasbourg, after the initial mistake of entering the data into as flight path angle instead of degrees, the crew failed to vigilantly validate the aircraft’s descent against other relevant forms of information, thus committing an error of omission. In the landing approach to Cali, the flight computer suggested the incorrect waypoint, ROMEO, and the crew committed a commission error by blindly following the automated suggestion and not adequately processing the information they received.

**Complacency.** Definitions of complacency include: “self-satisfaction that may result in non-vigilance based on an unjustified assumption of satisfactory system state,” and “a psychological state characterized by a low index of suspicion (Parasuraman and Manzey 2010)<sup>72</sup>”. Their research defines automation complacency as “poorer detection of system malfunctions under automation control compared with manual control”. This failure in achieving a fully coherent picture typically manifests itself under periods of high, multi-task work load, and constitutes an active diversion of attention from automation to other manual tasks. Although readily understood and accepted as undesirable, complacency presents a challenge in that complacent behavior may very seldom produce negative results since systems typically operate as expected. This can lead to failure of awareness and even possible acceptance of the behavior. In highly intensive and unforgiving systems like aircraft, however, all it can take is one unnoticed failure for there to be grave consequences. Complacency is closely related to automation bias as they both present manifestations of similar attentional issues. Most similarly, both automation complacency and automation bias can result in errors of omission. Complacency can result in this error from failure to appropriately monitor the automation itself due to diversion of attention, while automation bias results in failure to adequately monitor the system as a whole due to a bias that the automation will warn the operator if something goes wrong. All the case examples appear to exhibit complacent behavior to some degree. In Nagoya, the FO’s mistake of engaging the Go-Around mode could have been an innocent mistake, but both his failure to appropriately monitor the automation and fix the error along with the captain’s failure to ensure situation rectified lend themselves to complacent behavior. Both Strasbourg and Cali also show examples of incorrectly assumed satisfactory state of the system and automation although non-complacent behavior very possibly would have detected the mistakes in time.

**Mode Error.** Modes serves as a means through which automation can extend human capability by structuring complexity and presenting users with varying levels of control styles (i.e. “modes” of operation) according to Chappell et al. (1997)<sup>70</sup>. The authors note that glass cockpits have

---

<sup>70</sup> Chappell, Alan R, Edward G Crowther, Christine M Mitchell, and T Govindaraj. “The VNAV Tutor: Addressing a Mode Awareness Difficulty for Pilots of Glass Cockpit Aircraft.” *IEEE Transactions on Systems, Man and Cybernetics - Part A: Systems and Humans* , 1997: 372-385.

capitalized on the use of modes by giving pilots means to tailor the aircraft's automation to specific situations and preferences. Yet, as with most technology, new capabilities are closely paired with new pathways to failure. Specific to modes, a breakdown in coherence can occur when the human operator "loses track of which mode the device is in." Known as mode error, this breakdown results in a misinterpretation of the situation and unwanted system responses to given inputs (Sarter and Woods)<sup>73</sup>. Their research suggests that mode error occurs through a combination of "gaps and misconceptions" in operators' model of the automated systems and the failure of the automation interface to provide users with salient indications of its status and behavior. The Strasbourg accident clearly shows a crew committing mode error by failing to realize that they entered "33" into the descent rate mode rather than the desired descent angle mode. Had the crew maintained the proper awareness of the system's actual mode, they would have switched to the proper flight path angle mode without an issue and avoided their deadly error. Similarly at Nagoya, the aircraft responded appropriately given the Go-Around mode that was inadvertently commanded, yet the crew failed to understand the response of the aircraft and how to appropriately handle it, which ultimately led to the crash.

---

#### **COGNITIVE & PERCEPTUAL CHALLENGES**

Literature review and the analogy glass cockpit case exploration in this phase confirmed the need for further research into cognitive challenges in model-centric environments. These will be further explored in the next phase of research, along with strategies for mitigation.

In this phase, the research team has proposed six significant cognitive challenges to be applicable to IMCSE. Further discussion is found in German and Rhodes (2016)<sup>60</sup>.

These are:

#### **Automation Bias**

"The use of automation as a heuristic replacement for vigilant information seeking and processing," which may result in errors of commission and omission (Mosier and Skitka 1999)<sup>71</sup>

#### **Automation Complacency**

Assumed satisfactory state of system under automation control, resulting in poorer detection of system malfunctions compared to manual control (Parasuraman and Manzey 2010)<sup>72</sup>

#### **Mode Error**

---

<sup>71</sup> Mosier, Kathleen L, and Linda J Skitka. "Automation Use and Automation Bias." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Santa Monica: Human Factors and Ergonomics Society, 1999. 344-348

<sup>72</sup> Parasuraman, Raja, and Dietrich H Manzey. "Complacency and Bias in Human Use of Automation: An Attentional Integration." *Human Factors*, 2010: 381-410.

Failure to maintain awareness of the mode the system is presently in (Sater and Woods 1992)<sup>73</sup>

### **Attentional Tunneling**

Longer than optimal allocation of attention to a particular piece of information or task, resulting in neglect of other, more relevant concerns at hand (Wickens and Alexander 2009)<sup>74</sup>

### **Confirmation Bias**

“Seeking or interpreting of evidence in ways that are partial to existing beliefs, expectations, or a hypothesis at hand” (Nickerson 1998)<sup>75</sup>

### **Blind-Spot Bias**

Failure to acknowledge the existence of, or susceptibility to biases within oneself (Pronin et al. 2002)<sup>76</sup>

In addition to the cognitive challenges, perceptual challenges are also pertinent. It is necessary to not just take into account how information is cognitively processed, but also how information is perceived. The previous case examples have shown areas where glass cockpit technologies can contribute to cognitive failures, but additional research on the transition from analog to glass has revealed areas of perceptual failure. For example, a study by Wright and O’Hare (2015)<sup>77</sup> compared simulator flight performance between participants using tradition analog instruments and those using advanced glass cockpit displays, specifically comparing performance in loss of control events, and accuracy in maintaining altitude, airspeed, and heading. The results showed that the traditional cockpits actually resulted in better overall performance, corroborating with a separate study conducted by Hiremath et al. (2009)<sup>78</sup> which demonstrated that glass cockpit users had longer recovery times from unusual attitude situations than traditional cockpit users. In the Wright and O’Hare study, the pilot test participants unanimously rated the glass cockpit superior to the traditional displays. In their perception, the glass cockpit offered the “most awareness-enhancing, the least mentally demanding, and the easiest to interpret” display with the “fewest disliked features”. Despite this perceived superiority, the pilots actually performed worse with the display they preferred the most. This highlights a phenomenon known as ‘preference performance dissociation where users’ preferences do not line up with their performance. In the case of glass cockpits, Wright and O’Hare postulate that simply the use of

---

<sup>73</sup> Sarter, Nadine B, and David D Woods. “Mode Error in Supervisory Control of Automated Systems.” *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Santa Monica: Human Factors and Ergonomics Society, 1992. 26-29.

<sup>74</sup> Wickens, Christopher D, and Amy L Alexander. “Attentional Tunneling and Task management in Synthetic Vision Displays.” *The International Journal of Aviation Psychology*, 2009: 182-199.

<sup>75</sup> Nickerson, Raymond S. “Confirmation Bias: A Ubiquitous Phenomenon in Many Guises.” *Review of General Psychology*, 1998: 175-220.

<sup>76</sup> Pronin, Emily, Daniel Y Lin, and Lee Ross. “The Bias Blind Spot: Perceptions of Bias in Self Versus Others.” *Per Soc Psychol Bull*, 2002: 369-381.

<sup>77</sup> Wright, Stephen, and David O’Hare. “Can a glass cockpit display help (or hinder) performance of novices in simulated flight training?” *Applied Ergonomics* 47 (2015): 292-299.

<sup>78</sup> Hiremath, Vishal, Robert W Proctor, Richard O Fanjoy, Robert G Feyen, and John P Young. “Comparison of Pilot Recovery and Response Times in Two Types of Cockpits.” In *Human Interface and the Management of Information*, 766-775. San Diego: Springer Berlin Heidelberg, 2009

bright, highly contrasted colors results in the superior feedback as humans have been shown to prefer color as opposed to lack thereof. These findings strongly suggest that research is needed to understand more about the nature of human perception in model-centric environments in regard to what is most effective in the face of all the features technology can provide.

### **Mitigation Strategies**

Once the cognitive and perceptual challenges are better understood, appropriate strategies for mitigation may be formulated. Initially, heuristics may be derived from past cases and generalized knowledge, including from HSI and HCI knowledge. As the research continues into the next phase, the goal will be to mature the heuristics and formulate design principles and patterns that can be used in design and use of model-centric environments designed for effective human interactivity.

Four potential heuristics derived literature investigation in this phase are:

#### **Accountability**

Internalized responsibility for system performance that results in dedicated information seeking and decision making to ensure successful system performance while mitigating automation bias and complacency (Skitka et al. 2002; Mosier et al. 2000; )<sup>79, 68</sup>

#### **Transparent System**

System design that allows users to understand, follow, and predict the automation's behavior in order to reduce system surprises caused from mode error (Sarter et al. 1997; Besnard and Baxter 2006)<sup>80, 67</sup>

#### **Human-Centered Design**

Design focus towards achieving effective human integration with a system, rather than human adaptation to the system (Skitka et al. 2002)<sup>73</sup>

#### **Self-Awareness of Biases**

Educating users concerning potential biases that may influence decision making in order to preempt the formation and negative impact of biases (German and Rhodes 2016)<sup>60</sup>

---

## **DISCUSSION**

The use of analogy cases provides the means to investigate cognitive and perceptual considerations for interactive model-centric systems engineering practices and environments. Several analogy cases have been identified for further research. In this phase, the analogy case of pilots using glass cockpits has uncovered challenges that appear to be very relevant for the future of model-centric engineering. The use of advanced technology in cockpits manifests itself primarily through an increase in autonomy that not only changes the role of pilots, but also adds an additional component: manager of systems (Besnard and Baxter, 2006)<sup>81</sup>. While this

---

<sup>79</sup> Skitka, Linda J, Kathleen L Mosier, and Mark D Burdick. "Accountability and Automation Bias." *International Journal of Human-Computer Studies*, 2000: 701-717

<sup>80</sup> Sarter, Nadine B, David D Woods, and Charles E Billings. "Automation Surprises." In *Handbook of Human Factors & Ergonomics*, 2nd Ed, 1926-1943. Hoboken: John Wiley & Sons, 1997

<sup>81</sup> Besnard, Denis, and Gordon Baxter. "Cognitive Conflicts in Dynamic Systems." In *Structure for Dependability: Computer-Based Systems from an Interdisciplinary Perspective*, 107-124. London: Springer London, 2006.

technology has been successfully integrated into modern aviation, it also highlights the continued importance of considering the human interaction with technology in general, as specifically evidenced by disastrous examples. As revealed through the case research, cognitive coherence failures of automation bias, complacency, and mode error are possible challenges in model-centric environments. Mitigation heuristics of potential value include importance of accountability, transparency of systems, and human-centered design. Much can be learned by looking at historical cases, as well as adopting successful strategies from other fields. As experience grows with human-model interaction in model-centric engineering it is likely that unique challenges will be found.

---

#### **NEXT STEPS**

During the next phase, the team will continue investigating situations where humans interact with abstracted representations and models. Potential cases under consideration are a power plant control center, intelligent transportation models, supervisory control systems, and others. Literature review and gathering subject matter expert knowledge will be continued and expanded to identify heuristics, principles and mitigation strategies that are adoptable and adaptable from other fields such as Human-Computer Interaction (HCI) and Human Systems Integration (HSI). As knowledge is generated, the goal is to develop guidance practices that will have a positive impact on the effectiveness of human-model interaction. Collaboration with individual experts and professional societies is anticipated in accomplishing this goal.

---

## INTERACTIVE EPOCH-ERA ANALYSIS

Epoch-Era Analysis is a framework that supports narrative and computational scenario planning and analysis for both short run and long run futures. Because of the complex data that must be analyzed when extending EEA to large-scale problems, issues with cognition are introduced that may hamper decision-making. This motivates the need for extensions to EEA methods that overcome the computational and human cognition issues that may arise as a result. The Interactive Epoch-Era Analysis (IEEA) framework, comprised of 10 processes grouped in 6 modules, is introduced as a means for analyzing lifecycle uncertainty when designing systems for sustained value delivery. IEEA is proposed as an iterative framework for concept exploration that provides a means of applying EEA constructs while controlling growth in data scale and dimensionality. Further, IEEA leverages interactive visualization because prior visual analytics research has demonstrated that when performing exploratory analysis, like early-phase system concept selection, an analyst can gain deeper understanding of data which can lead to improved decision-making. It is hypothesized that the extension of interactive visualization to system design problems with lifecycle uncertainty may result in improved comprehension of the nature of underlying trades and improve a designer's ability to communicate their decision-making rationale. Application of IEEA to a case study for a multi-mission on-orbit servicing vehicle is provided to demonstrate key concepts and prototype interactive visualizations.

This project is performing exploratory development of interactive Epoch-Era Analysis, including human interface and reasoning considerations for epoch and era characterizations, as well as single and multi-epoch/era analyses.

---

## BACKGROUND

Epoch-Era Analysis (EEA) is an approach designed to clarify the effects of changing contexts over time on the perceived value of a system in a structured way<sup>82,83</sup>. The base unit of time in EEA is the epoch, which is defined as a time period of fixed needs and context in which the system exists. Epochs are represented using a set of epoch variables, which can be continuous or discrete values. These variables can be used to represent any exogenous uncertainty that might have an effect on the usage and perceived value of the system. Weather conditions, political scenarios, financial situations, operational plans, and the availability of other technologies are all potential epoch variables. Appropriate epoch variables for an analysis include key (i.e., impactful) exogenous uncertainty factors that will affect the perceived success of the system. A large set of epochs, differentiated using different enumerated levels of these variables, can then be assembled into eras, ordered sequences of duration-labeled epochs creating a description of a potential progression of contexts and needs over time. This approach provides an intuitive basis upon which to perform analysis of value delivery over time for systems under the effects of

---

<sup>82</sup> Ross AM. Managing unarticulated value: changeability in multi-attribute tradespace exploration. PhD in Engineering Systems. Cambridge, MA: MIT, 2006.

<sup>83</sup> Ross AM, Rhodes DH. Using natural value-centric time scales for conceptualizing system timelines through epoch-era analysis. INCOSE Int'l Symp 2008, Utrecht, the Netherlands, June 2008.

changing circumstances and operating conditions, an important step to take when evaluating large-scale engineering systems with long lifecycles.

Encapsulating potential short run uncertainty (i.e., what epoch will my system experience next?) and long run uncertainty (i.e. what potential sequences of epochs will my system experience in the future?) allows analysts and decision makers to develop dynamic strategies that can enable system value sustainment. Key challenges in application of EEA up to this point involve eliciting a potentially large number of relevant epochs and eras, conducting analysis across these epochs and eras, and extracting useful and actionable information from the analyses. Schaffner<sup>84</sup> showed that the number of potential eras to consider can grow very quickly, becoming computationally infeasible. As an example, an epoch space represented by 5 epoch variables, each with 3 levels, would result in  $3^5 = 243$  possible epochs. If the length of our eras is 10 epochs and each epoch can transition between any other epoch then the size of the potential era space would be  $243^{10} \sim 10^{24}$  eras. This means that for many problem formulations it is not feasible to evaluate systems across all or even a large fraction of potential eras.

EEA provides a powerful way of framing the problem of deciding among actionable alternatives given uncertainty, but comes at the expense of a potentially large and complex data set. The data that must be evaluated can be difficult to process, visualize and interpret by a decision maker. Traditional engineering and scientific visualization techniques may be inadequate for extracting insights from such data sets, but recent work in the area of visual analytics may prove helpful in mitigating these shortcomings. Visual analytics extends beyond traditional scientific visualization and focuses on extracting insights from data using interactive visual interfaces<sup>85</sup>. The research agenda in this area seeks to develop “the science of analytical reasoning facilitated by interactive visual interfaces”<sup>86</sup>. Good overviews of the state of current research on visual analytics are provided by Icke<sup>87</sup> and Keim<sup>88</sup>.

While there is much overlap, generally speaking researchers have been tackling the research from three angles: (1) data reduction and handling of large amounts of data; (2) specific types of visualizations that improve human cognition; and (3) methods to facilitate user interaction with data. Prior work has demonstrated promise for such capability and insight improvement when interactivity is added to tradespace exploration<sup>89</sup>. Supplying the decision-maker with immediate visual feedback on the consequences of their decisions could be enabled through simultaneous coordinated views of the design, performance and value spaces. Enabling users to interact with

---

<sup>84</sup> Schaffner MA. Designing systems for many possible futures: the RSC-based method for affordable concept selection (RMACS), with multi-era analysis. SM in Aeronautics and Astronautics. Cambridge, MA: MIT, 2014.

<sup>85</sup> Wong, P.C. and Thomas, J. "Visual Analytics". IEEE Computer Graphics and Applications, Vol. 24, Issue 5, 2004 Page(s): 20–21.

<sup>86</sup> Thomas, J. and Cook, K. "Illuminating the Path: The R&D Agenda for Visual Analytics National Visualization and Analytics Center," National Visualization and Analytics Center, 2005.

<sup>87</sup> Icke, I. and Sklar, E., "Visual Analytics: A Multi-faceted Overview," NSF Technical Report, April 16, 2009.

<sup>88</sup> Keim, D., Andrienko, G., Fekete, J., Gorg, C., Kohlhammer, J., and Melançon, G., "Visual analytics: Definition, process, and challenges," Information Visualization, pages 154–175. Springer, 2008.

<sup>89</sup> Ross AM, McManus HL, Rhodes DH, Hastings DE. A role for interactive tradespace exploration in multi-stakeholder negotiations. AIAA Space 2010. Anaheim, CA, September 2010.

their data through visual interfaces of this type is an area of active research<sup>90,91,92</sup>. Integrating interactive data visualization and advanced systems engineering methods is seen as key to the current research effort on IEEA.

---

## A FRAMEWORK FOR IEEA

IEEA leverages human-in-the-loop (HIL) interaction to manage challenges associated with the large amounts of data potentially generated in a study, as well as to improve sense-making of the results. By allowing the structured evaluation and visualization of many design alternatives across many different futures and potential lifecycle paths, this new approach enables the design of systems that can deliver sustained value under uncertainty.

### Extension of Prior EEA-based Methods

The framework described in this report is based on prior research on methods and processes for applying EEA constructs. The Responsive Systems Comparison (RSC) method, proposed by Ross et al.<sup>93,94</sup> as a prescriptive method for applying MATE and EEA, was developed to study system value sustainment through changeability. More recently, Schaffner<sup>95</sup> proposed the RSC-based Method for Affordable Concept Selection (RMACS) that expands the original seven processes of RSC to nine and explored the application of multi-attribute expense (MAE) to more effectively capture all resources expenditures required to realize a given system.

IEEA differs from both RSC and RMACS in that it strongly emphasizes iteration and human-in-the-loop (HIL) interaction throughout the process. Iteration is necessary because the analysis is inherently exploratory in nature. HIL interaction is necessary because the problem is not strictly deterministic or necessarily intended as a reliable prediction of system performance or future events. Often, there is both uncertainty and the potential for errors in assumptions or model implementation. This necessitates human judgment to make sense of the data, therefore this is not by its nature a problem that can be handed over completely to an automated optimization algorithm. Though some level of automated analysis could be beneficial as an aid to the user.

---

<sup>90</sup> Heer, J., and Schneiderman, B., "Interactive Dynamics for Visual Analysis," ACM Queue, 2012.

<sup>91</sup> Keim, D., "Information visualization and visual data mining," IEEE Trans. on Visualization and Computer Graphics, 8(1):1–8, 2002.

<sup>92</sup> Heer, J. and Agrawala, M., "Design considerations for collaborative visual analytics," Information Visualization, 7(1):49–62, 2007.

<sup>93</sup> Ross, A.M., McManus, H.L., Long, A., Richards, M.G., Rhodes, D.H., and Hastings, D.E., "Responsive Systems Comparison Method: Case Study in Assessing Future Designs in the Presence of Change," AIAA Space 2008, San Diego, CA, September 2008.

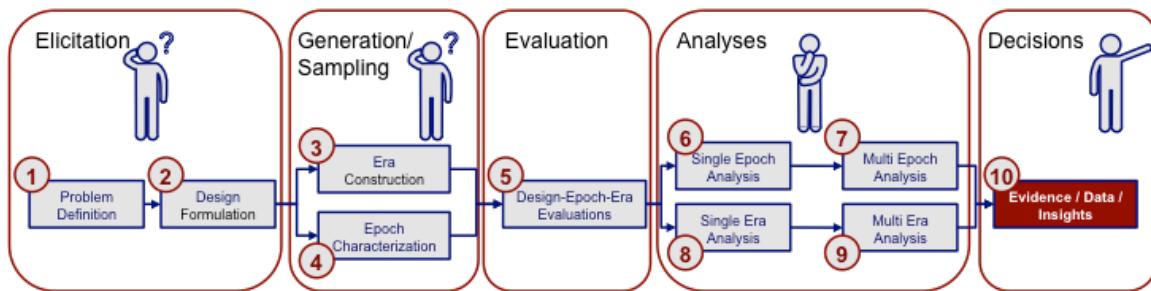
<sup>94</sup> Ross, A.M., McManus, H.L., Rhodes, D.H., Hastings, D.E., and Long, A.M., "Responsive Systems Comparison Method: Dynamic Insights into Designing a Satellite Radar System," AIAA Space 2009, Pasadena, CA, September 2009.

<sup>95</sup> Schaffner MA. Designing systems for many possible futures: the RSC-based method for affordable concept selection (RMACS), with multi-era analysis. SM in Aeronautics and Astronautics. Cambridge, MA: MIT, 2014.

## Description of IEEA Framework Modules

The purpose of IEEA, much like the purpose of RSC as described by Ross et al.<sup>96</sup>, is to “guide the...practitioner through the steps of determining how a system will deliver value, brainstorming solution concepts, identifying variances in contexts and needs (epochs) that may alter the perceived value delivered by the system concepts, evaluating key system trade-offs across varying epochs (eras) to be encountered by the system, and lastly developing strategies for how a designer might develop and transition a particular system concept through and in response to these varying epochs”. To that end, as shown in Figure 7, the IEEA framework is characterized by 10 individual processes that can be abstracted into six main modules:

1. **Elicitation** of relevant epoch and design variables (often through interview)
2. **Generation** of all epochs, eras and design tradespaces (often including enumeration)
3. **Sampling** of epochs and eras in which to evaluate design choices
4. **Evaluation** of designs in sampled subset of epochs and eras
5. **Analyses** of design choices in the previously evaluated epochs and eras
6. **Decisions** of final designs based on iterative evidence from previous modules.



**Figure 7. Interactive Epoch-Era Analysis processes and modules**

The modules and related activities are more clearly depicted in Figure 8.

---

<sup>96</sup> Ross, A.M., McManus, H.L., Long, A., Richards, M.G., Rhodes, D.H., and Hastings, D.E., "Responsive Systems Comparison Method: Case Study in Assessing Future Designs in the Presence of Change," AIAA Space 2008, San Diego, CA, September 2008.

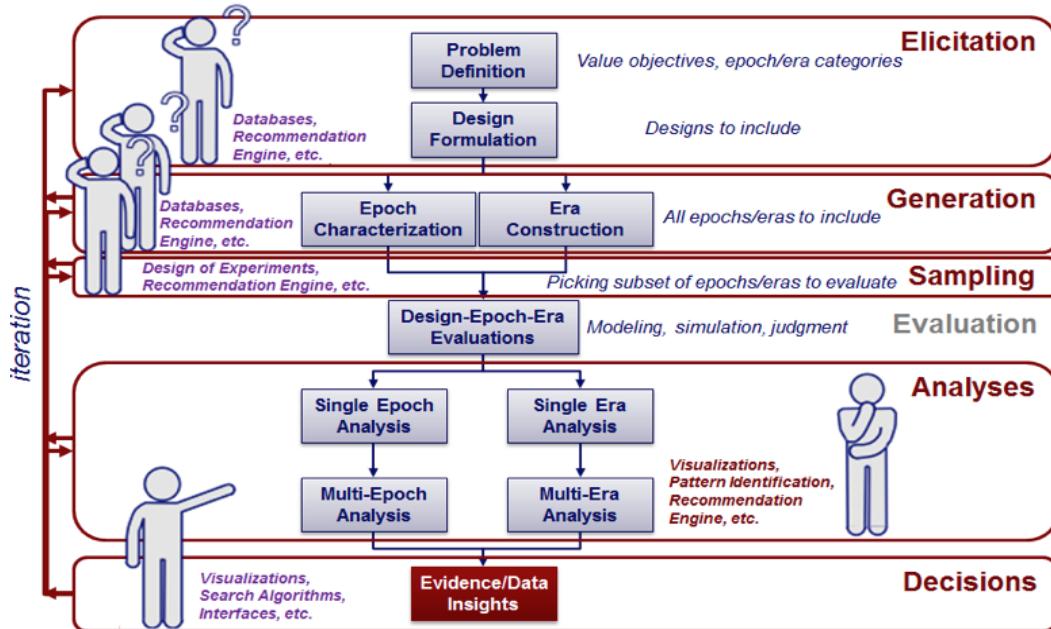


Figure 8. Interactive Epoch-Era Analysis modules<sup>97</sup>

While the sequence of these modules flows logically, IEEA is intended to be an iterative process where users can go back and change responses within earlier modules at any point to reflect what they have learned from later ones. The six modules are composed of the 10 processes described in further detail below, but depending on the nature of the study and the type and fidelity of information available to the analyst, it is not strictly required that each process step be applied. Many of the techniques discussed in Curry et al.<sup>98</sup> can be applied to augment and facilitate a practical implementation of the workflow. For example, OLAP techniques may be applied to improve data handling, and search algorithms may improve our ability to offer more informed recommendations to decision-makers during the epoch-era analysis process. Similarly, enhanced human interaction techniques and visualizations may aid in the analyses of the vast amounts of information required to reach an informed decision.

## Elicitation

### Process 1: Value-Driving Context Definition

The first process of the IEEA framework is to identify the overall problem that the system is intended to solve, why it's important, and who cares about the problem and solution. This requires the identification of important stakeholders, the resources available to them to solve the problem, and relevant exogenous uncertainties. Relevant exogenous uncertainties include those key contextual factors outside the system designer's control that may affect the problem or solution. The initial value proposition that describes how the system delivers value to the stakeholders should also be formed.

<sup>97</sup> Curry, M., and Ross, A.M., "Considerations for an Extended Framework for Interactive Epoch-Era Analysis," 13th Conference on Systems Engineering Research, Hoboken, NJ, March 2015.

<sup>98</sup> Curry, M., and Ross, A.M., "Considerations for an Extended Framework for Interactive Epoch-Era Analysis," 13th Conference on Systems Engineering Research, Hoboken, NJ, March 2015.

### **Process 2: Value-Driven Design Formulation**

In this process, the need and resource statements for each stakeholder as expressed by their individual objectives are defined. From the needs statements key system performance attributes (metrics) and multi-attribute utility (MAU) functions describing each stakeholder's preference for those attributes can be identified. From the resource statements expense attributes and corresponding multi-attribute expense (MAE) functions can be identified. Frequently, the only expense considered is the lifecycle cost of the system alternatives, but other expenses such as scarce materials, labor or time resources may also be considered. Finally, system solution concepts should be proposed based on past concepts or opinions from subject matter experts (SME). These concepts are then decomposed into the design variables of the system.

### **Generation / Sampling**

#### **Process 3: Epoch Characterization**

In the third process, uncertainties in context and needs are parameterized. These will be the epoch variables that are enumerated in process 5 to describe various possible future contexts and needs that may prevent system value delivery in spite of the designer's intentions. Uncertainties in stakeholder resource availability and usage preferences are also identified.

#### **Process 4: Era Construction**

This process constructs era timelines composed of multiple sequences of epochs with a set duration to create long-run descriptions of possible future scenarios a system may encounter. The activities in this process are in many ways analogous to those used in narrative or computational scenario planning. The future timelines can be constructed manually with the aid of expert opinion (narrative) or by implementing probabilistic models (computational), such as Monte Carlo simulation or Markov chain models, that define epoch transitions.

### **Evaluation**

#### **Process 5: Design-Epoch-Era Evaluation**

This process begins by enumerating discrete levels of design and epoch variables that can then be mapped to the performance and expense attributes of the system utilizing modeling and simulation. Stakeholders' utility and expense functions are then used to generate the MAU and the MAE for each design, within each epoch.

### **Analysis**

#### **Process 6: Single Epoch Analyses**

This process analyzes the MAU and MAE (cost) of alternatives within particular epochs, and is similar to what is traditionally referred to in literature as tradespace exploration. Feasible design solutions (those that are within required performance bounds) are visualized on an MAU vs. MAE scatterplot for any given epoch (time period of fixed operating context and stakeholder needs). Within-epoch metrics, such as yield (percent of all designs that are feasible), give an indication of the difficulty of a particular context and needs set for considered designs.

#### **Process 7: Multi-Epoch Analysis**

This process seeks to identify designs that value robust across changing contexts and needs by implementing short run passive and active strategies for value sustainment. After completing the traditional tradespace exploration activities of Process 6, in which the analyst compares potential designs within individual epochs, metrics are derived from measuring design properties across multiple (or all) epochs to provide insight into the impact of uncertainties on design candidates. In addition, resource usage can be analyzed to identify designs that are robust to the expense factors identified in Process 3 (e.g. decreasing budgets or labor availability).

#### **Process 8: Single-Era Analyses**

This process performs lifecycle path analysis to examine the impact of time-dependent uncertainties described by era timelines that are composed of unfolding sequences of future epochs that were created in process 4. By examining a particular series of epochs for a given length of time, decision-makers can identify potential strengths and weaknesses of a design and better understand the potential impact of path-dependent, long run strategies for value sustainment.

#### **Process 9: Multi-Era Analysis**

This process extends Process 8 by evaluating the dynamic properties of a system across many possible future eras, identifying patterns of strategies that enable value sustainment across uncertain long run scenarios.

### **Decision-Making**

#### **Process 10: Decisions and Knowledge Capture**

Decide on final design choices based on data generated and analyzed during previous processes. The purpose of this process is not only to capture the final decision that is made, but also the chain of evidence that led to that decision which can also be captured in a database or by some other knowledge management system. This information may prove useful to future studies by allowing the analysis of the rationale and specific assumptions that went into a decision.

---

#### **DEMONSTRATION CASE: SPACE TUG**

To demonstrate the application of the IEEA framework, a case study aimed at designing a multi-mission orbital transfer vehicle, or space tug, was selected. A space tug may be used for a variety of missions including observing, servicing or retrieving on-orbit spacecraft. The original case study described by McManus et al.<sup>99</sup> is, at first glance, a seemingly simple trade study, but despite the simplicity of the system model the analysis is actually nontrivial. Fitzgerald<sup>100</sup> expanded upon this case as a demonstration of his valuation approach for strategic changeability (VASC). The case study demonstrated here (in 10 processes) replicates the one by Fitzgerald. This provides

---

<sup>99</sup> McManus, H., and Schuman, T., "Understanding the Orbital Transfer Vehicle Trade Space," AIAA Space 2003, Long Beach, CA, September 2003.

<sup>100</sup> Fitzgerald, M.E., Managing Uncertainty in Systems with a Valuation Approach for Strategic Changeability, Master of Science Thesis, Aeronautics and Astronautics, MIT, June 2012.

for an interesting comparison since the application of IEEA leads to different insights that impact previous conclusions.

### **Process 1: Value-Driving Context Definition**

The first process defines the stakeholders, problem statement, exogenous uncertainties and the basic value proposition for the system. The problem statement, as described by Fitzgerald, is that the project sponsor would like to develop a space tug that can provide services to customers that collectively have eight different missions they need to perform. The space tug delivers value by meeting the demands of as many of those customers as possible for as long as possible. The ability to do so is driven not just by the nature of a given design alternative, but also by external factors like technology level that directly impact the performance attributes of the system.

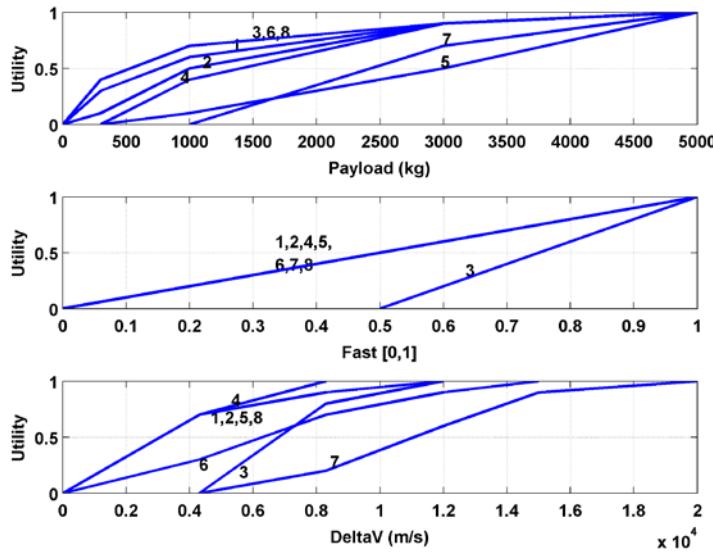
### **Process 2: Value-Driven Design Formulation**

The second process begins by defining the needs statements for all stakeholders, which become the attributes of system performance, along with utility functions describing each stakeholder's preference for each attribute. As shown in Table 3, for this study 8 different missions are defined, each with a different weighted preference for three system performance attributes: payload, speed and  $\Delta V$ .

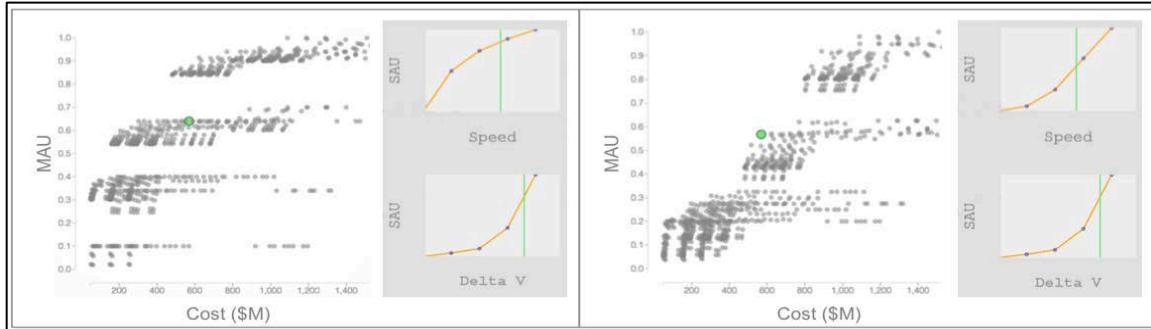
**Table 3. Summary of preferences for each mission**

	<b>Payload</b>	<b>Speed</b>	<b><math>\Delta V</math></b>
<b>(1) Baseline Mission</b>	0.3	0.1	0.6
<b>(2) Technology Demonstration</b>	0.7	0.1	0.2
<b>(3) GEO satellite rescue</b>	0.2	0.2	0.6
<b>(4) Satellite Deployment Assistance</b>	0.6	0.3	0.1
<b>(5) In-orbit refueling and maintenance</b>	0.75	0.05	0.2
<b>(6) Debris Collection</b>	0.2	0.05	0.75
<b>(7) All-purpose Military Mission</b>	0.4	0.2	0.4
<b>(8) Satellite Saboteur</b>	0.2	0.2	0.6

The value delivered by a given design alternative is different in each mission because they have different requirements (needs) and thus use different multi-attribute utility functions to calculate a measure of value based on the performance attributes of the design. For example, in the debris collector mission the single attribute weighting on speed is lower than in the rescue mission because that attribute is less important for that mission. The multi-attribute utility (MAU) function for each mission is developed from the weighted combination of the individual single attribute utilities (SAU). As shown in Figure 9, each performance attribute for each mission has a different mapping to a SAU value. For this case study each is a piecewise linear function, but SAU curves can take on more complex non-linear shapes depending on the needs of a particular stakeholder. For each design alternative, after computing the SAUs corresponding to the performance attributes, the MAU value can be computed as a weighted sum of the SAUs for each mission.



**Figure 9.** Single attribute utility (SAU) curves for the 8 different missions



**Figure 10.** Prototype Application for Interactive Visualization of MAU weightings showing (a) Initial tradespace and preferences over performance attributes; (b) Revised preferences and final tradespace<sup>101</sup>

Ricci, et al.<sup>102</sup> previously discussed a simplified version of the space tug case study as an example of how these SAU functions could be developed and weighted interactively using a HIL application as shown in Figure 10. A stakeholder may benefit from this type of interaction if the qualities of the system they believe to be valuable are not well articulated. This type of application is an example of how interactivity can often be a useful or necessary component of this process when eliciting stakeholder value statements.

### Process 3: Epoch Characterization

In process 3 the key contextual uncertainties are identified so that epoch variables can be characterized. In addition to different preference sets, value delivery for design alternatives in this case study is also affected by a single context variable, technology level, which has two levels, present or future. Technology level can directly impact the system performance attributes

<sup>101</sup> Ricci, N., Schaffner, M.A., Ross, A.M., Rhodes, D.H., and Fitzgerald, M.E., "Exploring Stakeholder Value Models Via Interactive Visualization," 12th Conference on Systems Engineering Research, Redondo Beach, CA, March 2014.

<sup>102</sup> ibid

through fuel efficiencies and vehicle mass fraction. It can also impact some transition costs when executing change options. A full-factorial design of all combinations of the 8 preference sets and 2 contexts results in 16 epochs as summarized in Table 4.

**Table 4. Summary of Epoch Variables**

Epoch Variable	Levels
Mission	1 through 8
Technology Level	Present or Future

#### **Process 4: Era Construction**

This process constructs era timelines composed of multiple sequences of epochs each with a set duration to create long-run descriptions of possible future scenarios a system may encounter. Simulating lifecycle performance in this way allows an analyst to evaluate path-dependent effects that may only arise when uncertainty is time-ordered. The activities in this process are in many ways analogous to those used in narrative or computational scenario planning. The future timelines can be constructed manually with the aid of expert opinion (narrative) or by implementing probabilistic models (computational), such as Monte Carlo simulation or Markov chain models, that define epoch transitions. For this case study, eras with a total length of 10 years, comprised of epochs uniformly distributed in duration between 1 to 12 months, were constructed according to the rules previously described by Fitzgerald<sup>103</sup>.

#### **Process 5: Design-Epoch-Era Evaluation**

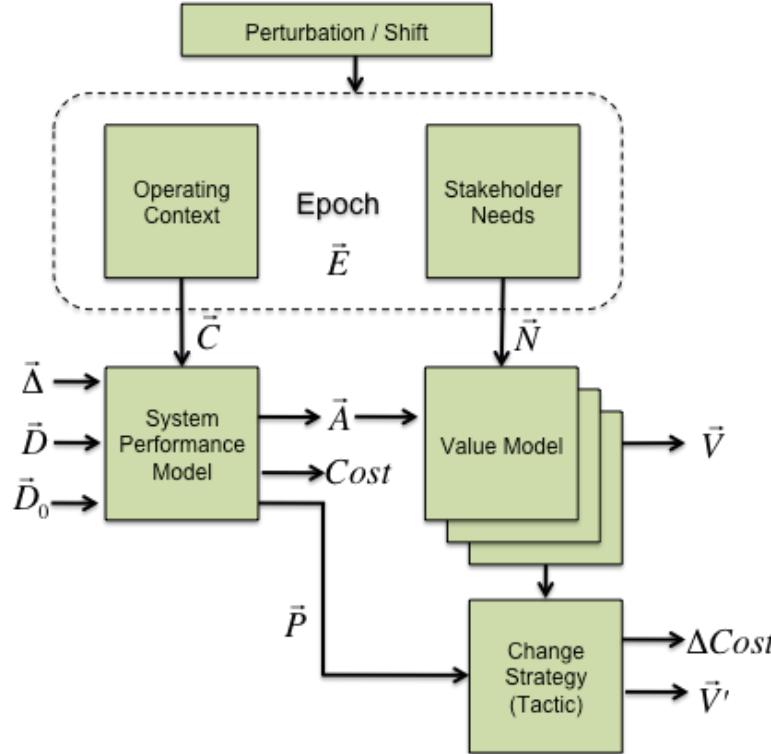
The first four processes defined the relevant elements of the models that will be evaluated in the fifth process. Figure 11 shows how the previously defined models are integrated to map design and epoch variables into stakeholder value (MAU) and expense (MAE). For this case study the only expense considered is the system cost. Fitzgerald<sup>104</sup> chose design variable levels as shown in Table 5 below. These are similar to the levels chosen by McManus et al.<sup>105</sup> except that a fourth design variable, design for changeability (DFC) level, is added. Both authors chose to use full-factorial experimental designs to enumerate the space of system performance attributes. This approach has been repeated here and the range of the evaluated performance attributes is shown in Table 6. A full-factorial enumeration of this space results in 432 designs, but not all designs are feasible. Due to compatibility constraints between propulsion type and fuel mass there are only 384 feasible designs.

---

<sup>103</sup> Fitzgerald, M.E., Managing Uncertainty in Systems with a Valuation Approach for Strategic Changeability, Master of Science Thesis, Aeronautics and Astronautics, MIT, June 2012.

<sup>104</sup> Fitzgerald, M.E., Managing Uncertainty in Systems with a Valuation Approach for Strategic Changeability, Master of Science Thesis, Aeronautics and Astronautics, MIT, June 2012.

<sup>105</sup> McManus, H., and Schuman, T., "Understanding the Orbital Transfer Vehicle Trade Space," AIAA Space 2003, Long Beach, CA, September 2003.



**Figure 11. Combinations of Design-Epoch-Era evaluated through performance, value, context and strategy models**

**Table 5. Design variable levels**

Design Variable	Levels
Capability	Low, Med, High, Extreme
Propulsion Type	BiProp, Cryo, Electric, Nuclear
Fuel Mass	30, 100, 300, 600, 1200, 3000, 10000, 30000, 50000
DFC Level	0, 1, 2

**Table 6. Performance attribute levels across all evaluated designs**

Performance Attributes	Range
Cost	[96.9, 3952.8]
Payload	[300, 5000]
Speed	[0, 1]
Delta V	[24.4, 41604.0]

Since the space tug model is relatively simple the computational time required to evaluate all design-context pairs is small. For more complex system models, fewer design levels or more sophisticated enumerating via experimental designs, such as central composite designs, may be required to lower the computational burden. Selecting the best way to enumerate the input space is another potential opportunity for improvement in this process through HIL interactivity,

but this has not yet been explored. After evaluating the performance space of available alternatives the value models developed in process 2 can now be applied to map the performance of each design-context pair into the value delivered in each of the 8 missions. This results in 6,144 design-epoch pairs (384 designs \* 2 context \* 8 preference sets) to consider. The following sections will discuss how these design-epoch pairs can be better understood through interactive visualization as demonstrated through various prototype applications.

### **Process 6: Single Epoch Analyses**

Single epoch analysis is comparable to what is often traditionally referred to in practice as tradespace exploration. Within a given epoch a scatter plot of cost (MAE) versus benefit (MAU) can be constructed that is fixed for short-run periods of stable context and needs (i.e., an epoch). Typically, a decision-maker wants to identify the frontier of Pareto optimal designs or, more generally, designs that are “close enough” to the Pareto front. Here the notion of “close enough” is operationalized through a metric called Fuzzy Pareto Number (FPN)<sup>106</sup>, based on the Fuzzy distance from a Pareto front<sup>107</sup> which is used to quantify the distance from the Pareto Front for each design in each epoch. FPN is a “within-epoch” metric and its value for a given design will change in different epochs. A decision-maker can gain insights regarding the difficulty of a particular context and needs by visualizing how points move in the design space as the epoch and FPN values change. Additional insights may come from interactively filtering the design, performance or value variables. This can be performed with the aid of the filtering application shown in Figure 12 that allows the decision-maker to interact with their data to identify designs and epochs of interest. It also allows them to assign any of the defined variables to the radius, color or x-y location of the points in the scatter plot to explore the data in four dimensions and better comprehend the behavior of the designs.

---

<sup>106</sup> Fitzgerald, M.E. and Ross, A.M., "Mitigating Contextual Uncertainties with Valuable Changeability Analysis in the Multi-Epoch Domain," 6th Annual IEEE Systems Conference, Vancouver, Canada, March 2012.

<sup>107</sup> Smaling, R., "Fuzzy Pareto Frontiers in Multidisciplinary System Architecture Analysis", 10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, Albany, NY, September 2004.

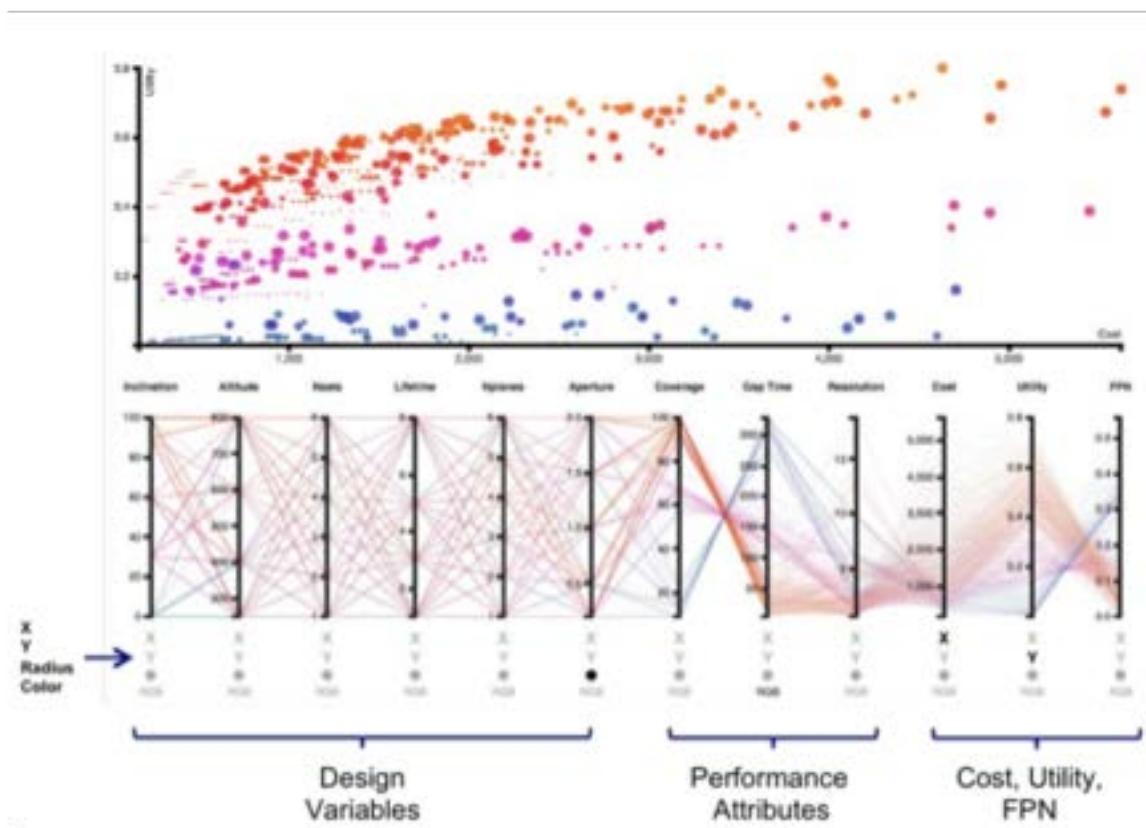


Figure 12. Interactive Filtering Application

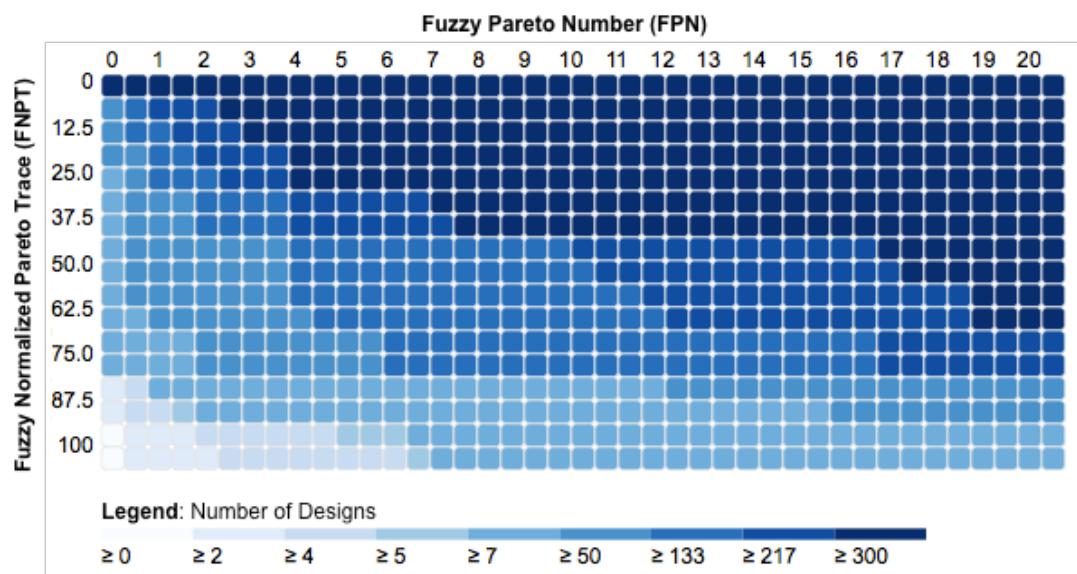


Figure 13. Interactive heatmap visualization



Figure 14. Interactive Filtering Application implementing OLAP.

### Process 7: Multi-Epoch Analysis

The activities of process 7 allow a decision-maker to gain deeper insights by evaluating metrics between and across epochs to gauge the impact of uncertainties on system value. This includes the evaluation of short run passive and active strategies for achieving value sustainment such that systems can maintain value delivery across different missions or changing contexts. A system that is passively robust is insensitive to changing conditions and continues to deliver acceptable value. Alternatively, a system that suffers deterioration in value due to evolving conditions may benefit from the use of change options that make them flexible, adaptable or resilient.

#### Evaluating passive strategies for value sustainment (Robustness)

Ideally a design candidate would be Pareto optimal in each of the 16 defined epochs and be within the required cost and performance constraints set by each stakeholder. This is often unrealistic, however, so a decision-maker may be required to settle for a design that is close enough to the Pareto front across most epochs. As was the case in process 6, “close enough” is operationalized through the FPN metric, but this analysis also must define a metric that captures the frequency at which a particular design meets a threshold FPN across epochs. To accomplish

this the Fuzzy Normalized Pareto Trace (FNPT) metric<sup>108,109</sup> is defined as the percentage of epochs in which a given design appears within a range from the Pareto front defined by the analyst. Applying these two metrics a decision-maker can set a threshold FPN and evaluate how frequently a design appears close to the Pareto front across all epochs. Assuming no designs are Pareto optimal in every epoch, a decision-maker can choose to relax the acceptable distance from the Pareto front by increasing the FPN threshold or accept a lower FNPT indicating decreased Pareto efficiency of the design in some epochs.

In past studies the trade-off between FPN and FNPT has been a very manual process that may benefit from implementation in an interactive application. The single-epoch analysis application shown in Figure 12 can be impractical for this analysis if the number of epochs and/or designs is large. Binned aggregation techniques as discussed in Curry et al.<sup>110</sup> can be applied to overcome these types of issues. The interactive heatmap visualization in Figure 13 shows the tradeoff between FPN and FNPT using color to encode the number of designs that satisfy the threshold at each level. Clicking on any square in the heatmap brings up a separate list of the designs that meet the cutoff. If an analyst would like to concurrently examine the impact of various FPN and FNPT trades on design and performance variables a more complex visualization can be implemented using OLAP to handle issues that arise with more data dimensions and an increasingly larger data set that must be manipulated in real-time. As an example, the interactive visualizations shown in Figure 14 applies OLAP, multiple coordinated views and binned aggregation to allow trade-offs between Pareto efficiency (FPN) and frequency of acceptable epoch performance (FNPT). This application also allows a decision-maker to determine not just the percentage of acceptable epochs, but also which epochs are most difficult for candidate designs. This is an insight not previously available or discussed in prior applications of multi-epoch analysis for this case study. These types of previously undiscovered relationships and patterns within the dataset may be useful for identifying “problem epochs” or when determining cases where it might be more appropriate to build a combination of systems to satisfy all possible future epochs.

Applying this approach to the space tug case study it can be shown that none of the enumerated designs are Pareto optimal (FPN=0%) all of the time (FNPT=100%). Depending on the preferences of the decision-maker, using the interactive filtering application they could now choose to relax the FPN or the FNPT constraint to identify acceptable design compromises. Holding the requirement on Pareto efficiency (FPN=0%) constant and relaxing the requirement on FNPT it can be determined that 3 designs are Pareto optimal in 14 of 16 (FNPT=87.5%) of enumerated epochs as shown in Table 7. Alternatively, a decision-maker may be more willing to relax their requirement on Pareto efficiency (FPN=1%) and require that this threshold be met in all epochs

---

<sup>108</sup> Ross, A.M., McManus, H.L., Long, A., Richards, M.G., Rhodes, D.H., and Hastings, D.E., "Responsive Systems Comparison Method: Case Study in Assessing Future Designs in the Presence of Change," AIAA Space 2008, San Diego, CA, September 2008.

<sup>109</sup> Fitzgerald, M.E., Managing Uncertainty in Systems with a Valuation Approach for Strategic Changeability, Master of Science Thesis, Aeronautics and Astronautics, MIT, June 2012.

<sup>110</sup> Curry, M, and Ross, A.M., "Considerations for an Extended Framework for Interactive Epoch-Era Analysis," 13th Conference on Systems Engineering Research, Hoboken, NJ, March 2015.

(FNPT=100%) as shown in Table 8. Finally, if a decision-maker was willing to relax both the FPN and FNPT requirement more designs remain after interactively filtering as shown in Table 9. Comparing the designs identified here to those previously identified as passively robust by Fitzgerald<sup>111</sup>, shown in Table 10, it can be seen that only 2 of the 5 designs are the same, designs 128 and 191. Notably, these two designs were shown by Fitzgerald to be among the best performing of his identified designs in subsequent multi-epoch and multi-era analysis.

**Table 7. 3 designs are within 0.0% (FPN) of Pareto optimal in 87.5% (FNPT) of enumerated epochs**

Design#	Cost (\$M)	Capability	Engine Type	Propellant Mass (kg)	DFC Level
63	900	Medium	Nuclear	10000	0
95	1540	High	Nuclear	10000	0
128	3020	Extreme	Nuclear	30000	0

**Table 8. 3 designs are within 1.0% (FPN) of Pareto optimal in 100% (FNPT) of enumerated epochs**

Design#	Cost (\$M)	Capability	Engine Type	Propellant Mass (kg)	DFC Level
63	900	Medium	Nuclear	10000	0
128	3020	Extreme	Nuclear	30000	0
191	980	Medium	Nuclear	10000	1

**Table 9. 5 are within 1.0% (FPN) of Pareto optimal in 87.5% (FNPT) of enumerated epochs**

Design#	Cost (\$M)	Capability	Engine Type	Propellant Mass (kg)	DFC Level
30	382	Low	Nuclear	3000	0
63	900	Medium	Nuclear	10000	0
95	1540	High	Nuclear	10000	0
128	3020	Extreme	Nuclear	30000	0
191	980	Medium	Nuclear	10000	1

**Table 10. Passively robust designs identified by Fitzgerald<sup>19</sup>**

Design#	Cost (\$M)	Capability	Engine Type	Propellant Mass (kg)	DFC Level
1	97	Low	Biprop	30	0
29	306	Low	Nuke	1200	0
47	628	Low	Cryo	10000	0
128	3020	Extreme	Nuclear	30000	0
191	980	Medium	Nuclear	10000	1

### Evaluating active strategies for value sustainment (Changeability)

---

<sup>111</sup> Fitzgerald, M.E., Managing Uncertainty in Systems with a Valuation Approach for Strategic Changeability, Master of Science Thesis, Aeronautics and Astronautics, MIT, June 2012.

The multi-epoch analysis metrics described above show how value sustainment through design robustness can be evaluated for this case study. A design that is not robust to changes in needs or context, however, may still be able to sustain value through the use of design change options. A system that is equipped with a design feature, or option, that allows it to change its state may do so to restore value if a future epoch is encountered that causes a loss of value. Typically, these change options are built into a design at the beginning of its life for an additional cost, but not used unless a particular future unfolds. There may also be an associated cost in money, time or other resources to execute the option. For this case study there are six change options defined as shown in Table 11.

**Table 11. Available Change Mechanisms**

No.	Change Option	Effect	DFC Level
1	Engine Swap	Biprop/Cryo swap	0
2	Fuel Tank Swap	Change fuel mass	0
3	Engine Swap (reduced cost)	Biprop/Cryo swap	1 or 2
4	Fuel Tank Swap (reduced cost)	Change fuel mass	1 or 2
5	Change Capability	Change Capability	1 or 2
6	Refuel in Orbit	Change fuel mass (no redesign)	2

A logical series of questions that a decision-maker would want to answer next are, “Which of these options should be implemented to allow a candidate system to sustain value, what metrics allow that to be assessed and are the options worth the cost?” A useful metric proposed by Ross<sup>112</sup> is Filtered Outdegree (FOD), which represents the number of change paths out of a design to various target designs given a set of filtering constraints on resource usage such as execution time and expense (cost). Since FOD only captures the number of change paths, not necessarily whether they are valuable or useful, several follow-on works attempted to define metrics for various versions of Value-Weighted Filtered Outdegree (VWFO)<sup>113</sup> to assess the utility gain that could be achieved through execution of various change options. Focusing more on long run strategies, Schaffner<sup>114</sup> provides a convincing argument that FOD may be a more appropriate metric than any of the proposed VWFO metrics since in many cases it may be beneficial to execute a change to affect a short-term loss of utility in order to achieve a longer-term net gain. Schaffner<sup>115</sup> further notes that most past applications of FOD have focused on single arc change paths to determine the number of target end states that an original design can achieve via execution of a change mechanism. His proposed metric for Fully Accessible Filtered Outdegree

---

<sup>112</sup> Ross AM. Managing unarticulated value: changeability in multi-attribute tradespace exploration. PhD in Engineering Systems. Cambridge, MA: MIT, 2006.

<sup>113</sup> Viscito, L., and Ross, A.M., "Quantifying Flexibility in Tradespace Exploration: Value-Weighted Filtered Outdegree," AIAA Space 2009, Pasadena, CA, September 2009.

<sup>114</sup> Schaffner MA. Designing systems for many possible futures: the RSC-based method for affordable concept selection (RMACS), with multi-era analysis. SM in Aeronautics and Astronautics. Cambridge, MA: MIT, 2014.

<sup>115</sup> ibid

(FAFO), which also captures additional end states that can be reached through multiple change mechanisms, is adopted for the research presented here on IEEA.

While FAFO can provide a proxy metric for desirable system behavior the strategy for how and when to execute it is equally as important when assessing the potential for a system to sustain value delivery. For a system that has one or more change mechanisms, the decision as to which option to execute or whether an option should be executed at all depends on the change strategy assumed. Fitzgerald<sup>116</sup> and Schaffner<sup>117</sup> describe several epoch-level change strategies that may be considered such as “maximize utility” or “maximize efficiency”. For the “maximize utility” strategy, as the name implies, it is assumed that a system will execute the change option or options that allow it to achieve the highest utility end state in a given epoch. Similarly, the “maximize efficiency” strategy will execute whichever options allow it to maximize Pareto efficiency and get closest to the Pareto front even if that results in a lower utility. The two authors also proposed other strategies, such as “Survive” and “Maximize Profit”, but it is possible that additional as yet undefined strategies may exist.

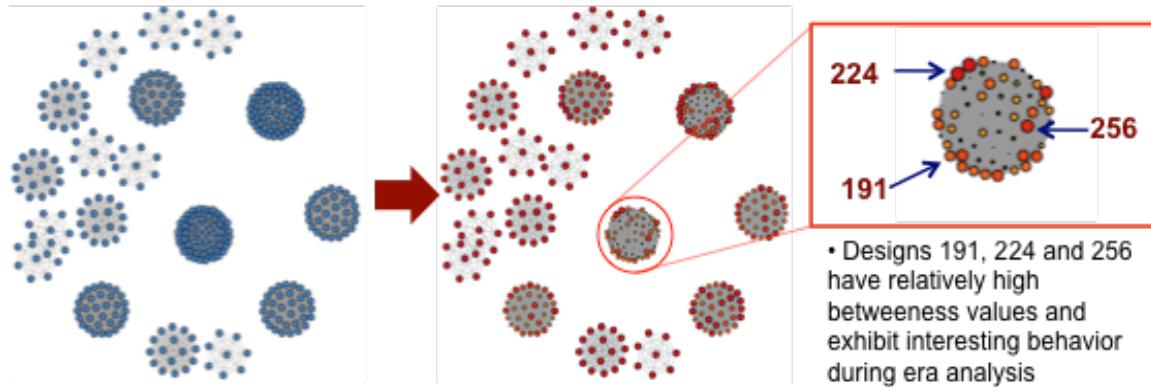
Though adjacency matrices have previously been used to visualize design change networks, interactive force directed graphs, as shown in Figure 15, might facilitate deeper insights from changeability analysis. For instance, the existence of clusters of designs related to one another through change mechanisms is apparent even without applying clustering algorithms such as the Louvain community detection method. Also, compared to the static adjacency matrix representations shown by Fitzgerald<sup>118</sup>, interactive visualization allows the represented network to be explored using dynamic data filtering and by changing the variable assignments of the visual elements. Using controls and filters that allow an analyst to assign node color, node radius, link width, link length and link color to different data dimensions an analyst can more readily identify unexpected results. In the example shown in Figure 15 the visualization on the right shows the result from interactively assigning network centrality metrics for betweenness to both node color and radius. Doing so provides immediate visual feedback that designs, even within the same cluster, are not necessarily changeable in the same way in terms of available end states and costs to reach them.

---

<sup>116</sup> Fitzgerald, M.E., Managing Uncertainty in Systems with a Valuation Approach for Strategic Changeability, Master of Science Thesis, Aeronautics and Astronautics, MIT, June 2012.

<sup>117</sup> Schaffner MA. Designing systems for many possible futures: the RSC-based method for affordable concept selection (RMACS), with multi-era analysis. SM in Aeronautics and Astronautics. Cambridge, MA: MIT, 2014.

<sup>118</sup> Fitzgerald, M.E., Managing Uncertainty in Systems with a Valuation Approach for Strategic Changeability, Master of Science Thesis, Aeronautics and Astronautics, MIT, June 2012.



**Figure 15. (a) Interactive force-directed graphs of change mechanisms between designs; (b) Betweenness centrality metric assigned to node radius**

The value of a change mechanism to a given design will vary based on the strategy assumed. To evaluate the benefits of various strategies and change option pairs for a set of designs, two metrics, effective FPN (eFPN) and effective FNPT (eFNPT), can be used<sup>119</sup>. The eFPN and eFNPT metrics can be evaluated for a particular design across all epochs for each strategy. If the change strategy dictates that an original design changes to a particular target design in a given epoch, then the target design is evaluated. If the change strategy dictates that a starting design does not change in that epoch, then that starting design is evaluated<sup>120</sup>. Assuming a usage strategy, we could now evaluate the valuable changeability of enumerated designs via the interactive application shown in Figure 15. Though not shown in this example, rule removal studies as proposed by Fitzgerald<sup>121</sup>, could also be implemented in this interactive application to allow a decision-maker to develop intuition regarding the value to a system of having a particular change options. These have been omitted from the current discussion for brevity.

### Summary of Multi-Epoch Analysis

The analyses outlined in this section provide a way for decision-makers to interactively evaluate the performance of multiple design alternatives across multiple futures. This creates opportunities for new insights at the expense of a potentially large and complex data set that can be difficult to make sense of even for this simplified case study. The application of an interactive framework allows the user to visualize and engage with the data in new ways that may facilitate improved comprehension and decision-making. The insights that can be extracted from this approach allow the decision-maker to understand the characteristics of designs that can sustain value in all possible futures through passive robustness or active changeability. Note that while it has been demonstrated here as two separate analyses it would be desirable, though not yet

<sup>119</sup> Fitzgerald, M.E., Managing Uncertainty in Systems with a Valuation Approach for Strategic Changeability, Master of Science Thesis, Aeronautics and Astronautics, MIT, June 2012.

<sup>120</sup> Schaffner MA. Designing systems for many possible futures: the RSC-based method for affordable concept selection (RMACS), with multi-era analysis. SM in Aeronautics and Astronautics. Cambridge, MA: MIT, 2014.

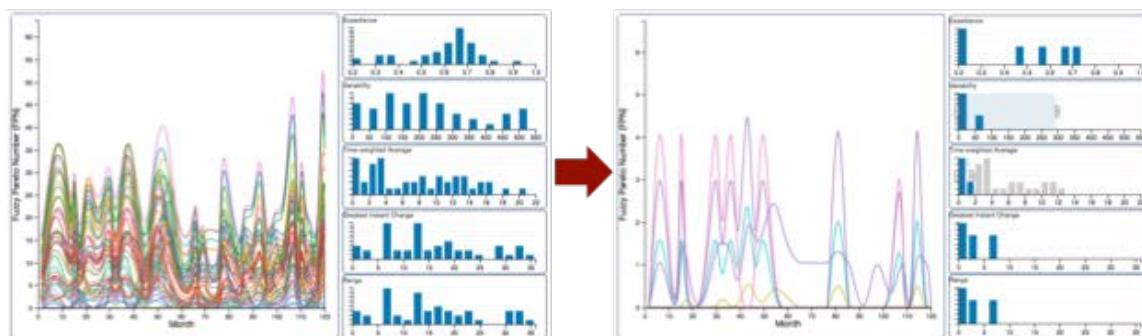
<sup>121</sup> Fitzgerald, M.E., Managing Uncertainty in Systems with a Valuation Approach for Strategic Changeability, Master of Science Thesis, Aeronautics and Astronautics, MIT, June 2012.

implemented, to concurrently evaluate robustness and changeability of designs. Designs that are moderately effective at both may be preferable to superior performance in only one.

### Process 8: Single-Era Analyses

Epoch-analysis is focused on the evaluation of short run passive and active strategies for achieving value sustainment. In contrast, era-analysis focuses on long run sustainment of system value delivery across different missions or changing contexts. This process examines the time-dependent effects of an unfolding sequence of future epochs created in Process 4. By examining a particular series of epochs for a given length of time, decision-makers can identify potential strengths and weaknesses of a design and better understand the potential impact of path-dependent, long run strategies for value sustainment.

For many system design applications, subject matter experts may identify eras from one or more likely narratives that may play out. When analyzing any one of those eras a decision-maker would then want to identify the right combination of inherent robustness, changeability and operational strategy that allow a system to meet a specified performance threshold across all future time steps. As an example, assume a design is desired that remains within a given distance of the Pareto front (e.g. FPN close to zero) across all time. A plot of the FPN versus time can reveal how each candidate design performs, but it is difficult to compare performance between designs across the era. Previous applications of era analysis<sup>122,123</sup> used variations of time-weighted average performance across an era to compare designs to one another. They also focused on utility, rather than FPN, as their metric of interest which is not necessarily appropriate since utility values are not comparable in different epochs. Schaffner<sup>124</sup> identified several additional metrics that can be applied to evaluate additional characteristics of performance across an era including expedience, variability, time-weighted average, greatest instant rise/fall, and range. These additional metrics provide an improved ability to describe era performance at the expense of increased information that a decision-maker must consider when selecting a design.



**Figure 16. Interactive application showing candidate designs across a single era (a) before filtering; and (b) after**

<sup>122</sup> Fitzgerald, M.E., Managing Uncertainty in Systems with a Valuation Approach for Strategic Changeability, Master of Science Thesis, Aeronautics and Astronautics, MIT, June 2012.

<sup>123</sup> Richards, M.G., "Multi-Attribute Tradespace Exploration for Survivability," Doctor of Philosophy Dissertation, Engineering Systems Division, MIT, June 2009.

<sup>124</sup> Schaffner MA. Designing systems for many possible futures: the RSC-based method for affordable concept selection (RMACS), with multi-era analysis. SM in Aeronautics and Astronautics. Cambridge, MA: MIT, 2014.

As demonstrated for previous IEEA processes, decision-making in single-era analysis can also benefit from the application of techniques such as multiple coordinated views, interactive filtering and OLAP. An example application, shown in Figure 16, shows the FPN performance of all designs across time for the specified era. The five histograms to the right display aggregated data on the performance of the candidate designs for the era metrics identified by Schaffner<sup>125</sup>. Interactive filtering on those metrics allows a decision-maker to rapidly identify interesting designs based on their individual preferences for average performance and stability of performance across time. It also allows them a way to better comprehend design behavior that has not previously been demonstrated.

### Process 9: Multi-Era Analysis

This process extends Process 8 by evaluating the dynamic properties of systems across many possible future eras, identifying patterns of strategies that enable value sustainment across uncertain long run scenarios. When looking at only a single era it is possible to compare how individual designs perform relative to one another using the era metrics previously discussed that capture temporal aspects of value delivery. This is not practical when analyzing many possible eras. In fact, it has been previously shown that it would be impossible to characterize the entire era-space<sup>126,127</sup>. The goal then for multi-era analysis is focused more on understanding the aggregate behavior of designs given different long-run strategies for operating a system. Specifically, it is useful to better understand any possible path dependencies that may arise due to either external perturbations/shifts or the application of operational strategies that define usage rules for available design change options.

In past research on path-dependency analysis for multiple eras, the progression of epochs within an era has been modeled as a directed acyclic graph<sup>128</sup> or tree of events<sup>129</sup>. This is similar to how path-dependency analysis is conducted in programs that analyze strategy games such as tic-tac-toe where the decision tree is searched using variants of the minimax algorithm to determine the best move at each step. In more complex games, such as chess, partial tree searches are typically required to keep the problem computationally tractable and the decision approaches optimal with increasing depth of the search through the tree<sup>130</sup>. HIL interaction, however, can be leveraged to enable the decision tree to be searched more efficiently. The benefits of HIL interaction has been demonstrated on related path-analysis problems such as the traveling

---

<sup>125</sup> ibid

<sup>126</sup> ibid

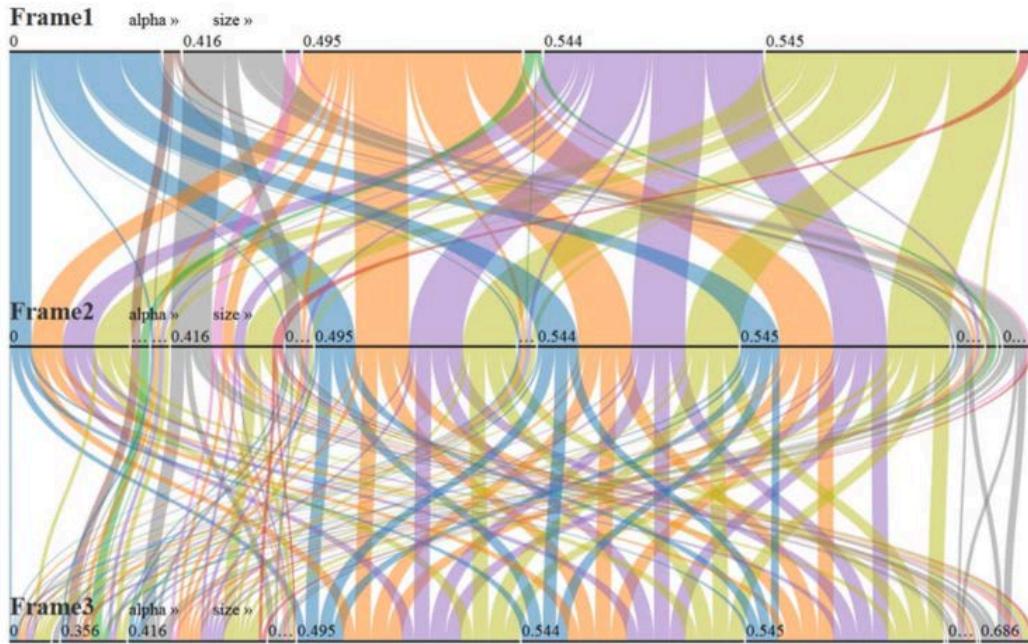
<sup>127</sup> Curry, M, and Ross, A.M., "Considerations for an Extended Framework for Interactive Epoch-Era Analysis," 13th Conference on Systems Engineering Research, Hoboken, NJ, March 2015.

<sup>128</sup> Note that eras could also be modeled using a directed cyclic graph. For instance if Markov Chain Monte Carlo (MCMC) methods were applied and transitions between a finite number of defined epochs were modeled probabilistically as was attempted by Fulcoly, D.O., Ross, A.M., and Rhodes, D.H., "Evaluating System Change Options and Timing using the Epoch Syncopation Framework," 10th Conference on Systems Engineering Research, St. Louis, MO, March 2012.

<sup>129</sup> Schaffner MA. Designing systems for many possible futures: the RSC-based method for affordable concept selection (RMACS), with multi-era analysis. SM in Aeronautics and Astronautics. Cambridge, MA: MIT, 2014.

<sup>130</sup> The exception is so called “pathological” game trees in Nau, D.S., “Pathology on Game Trees Revisited, and an Alternative to Minimaxing,” Artificial Intelligence, 221-244, 1983.

salesman problem<sup>131</sup> and “human-machine” chess matches that demonstrated a human player, coupled with chess software, can fairly consistently beat computer-only players<sup>132</sup>. This suggests that multi-era path analysis could also benefit from the right combination of interactive applications that leverage the experience of subject matter experts (SME) to identify beneficial or detrimental path-dependencies within eras.



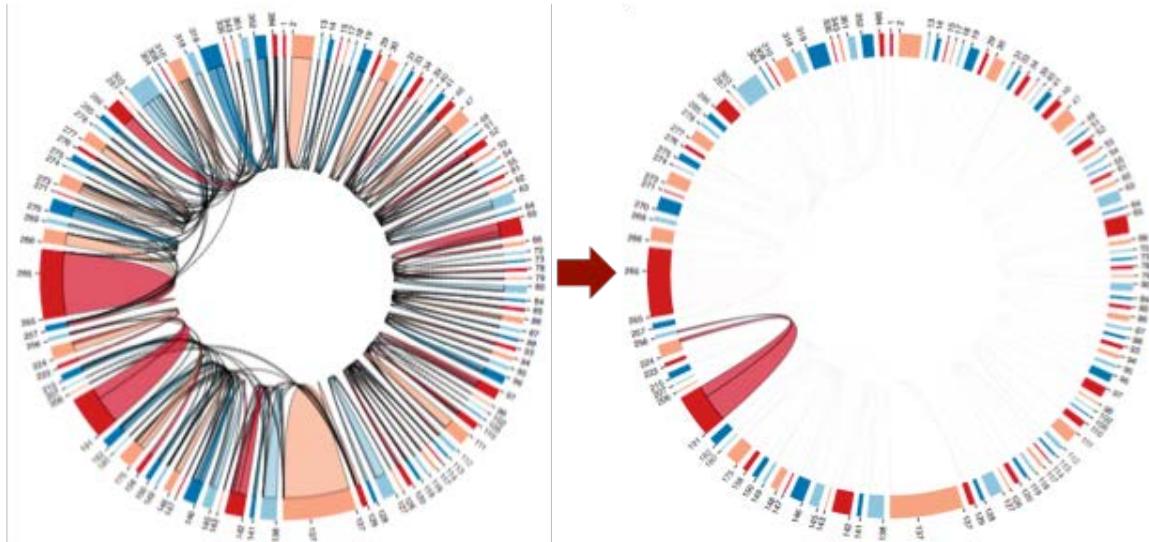
**Figure 17. Visualization showing eras visualized using parallel sets<sup>133</sup>**

---

<sup>131</sup> Krolak, P., Felts, W., and Marble, G., “A Man-Machine Approach Toward Solving the Traveling Salesman Problem,” Communications of the ACM, Volume 14, Number 5, May 1971.

<sup>132</sup> Thompson, C., “Smarter Than You Think : How Technology Is Changing Our Minds For The Better”. The Penguin Press, 2013.

<sup>133</sup> Schaffner MA. Designing systems for many possible futures: the RSC-based method for affordable concept selection (RMACS), with multi-era analysis. SM in Aeronautics and Astronautics. Cambridge, MA: MIT, 2014.



**Figure 18. Interactive chord diagram used to visualize design change behavior.**

As an example of how interactive multi-era analysis can generate insights into system behavior and long-run value delivery we can look at path dependencies that arise due to changeability usage for the space tug case study. Prior multi-era research<sup>134</sup> has examined the use of a parallel sets to visualize the proportion of design occurrences within the time duration of a particular epoch as shown in Figure 17. This type of visualization is useful in visualizing the temporal aspects of change across an era, but can become cumbersome for analysis across many long eras with many epochs. A new visualization, an interactive chord diagram, is introduced here as one possible way of representing aggregate change behavior in a more compact form. As shown in Figure 18, the chord diagram can be used to represent the proportion of the time that a source design executes a change option to reach various target designs across multiple eras. All the designs that use change options to sustain value are enumerated around the circumference of the diagram and quadratic Bézier curves show the proportion of each source design changing to each target. The source and target arcs represent mirrored subsets of aggregate change behavior. Detailed analysis using this visualization allows an analyst to quickly identify designs that rely on changeability (rather than robustness) to maintain value and which options and end-state designs that are frequently used for various strategies.

The multi-era change path dependency analysis shows that while the network of changes (due to execution of options) that manifest during the space tug analysis are complex, interaction can allow specific insights to be extracted. For example, from the interactive chord diagram visualization shown in Figure 18 we see that only a small fraction of designs, 109 out of 384 (28%), actually use changeability to maintain value when implementing the multi-era maximize efficiency strategy. By hovering over a specific design we can gain more detailed information about its behavior across eras. In this example, design 191 is shown to exist in a change “limit cycle” with designs 224 and 256. When executing a change option design 191 will change 29% of the time into design 224 and 71% of the time into design 256. We can also observe from this

<sup>134</sup> Schaffner MA. Designing systems for many possible futures: the RSC-based method for affordable concept selection (RMACS), with multi-era analysis. SM in Aeronautics and Astronautics. Cambridge, MA: MIT, 2014.

interactive visualization that design 191 is the target end state rather than the source design 8 times less frequently. When it is the end state, design 224 is the source 66% of the time and 256 the remainder of the time. That these 3 designs never transition to any of the other designs within their reachable network (family/community), which contains 32 total designs, highlights their relative importance in this strategy. This behavior can be shown to vary significantly with the strategy assumed. This emphasizes the point that system value delivery is sensitive not only to variables related to design, change options and context, but also to operational characteristics that must be considered.

### **Process 10: Decisions and Knowledge Capture**

The purpose of this process is not only to capture the final decision that is made, but also the chain of evidence that led to that decision which can be captured in a database or other knowledge management system. This information may prove useful to future studies by allowing post-hoc analysis of the rationale and specific assumptions that went into a decision. Though not demonstrated for this case study, the capability to capture and store key information about the reasoning behind a decision is an advantage of implementing IEEA in an integrated application such as the web-based application demonstrated here.

### **Summary of Case**

The research presented here introduces the Interactive Epoch-Era Analysis (IEEA) framework which provides a means for analyzing lifecycle uncertainty when designing systems for sustained value delivery. Application of IEEA to a case study for a multi-mission on-orbit servicing vehicle demonstrates key concepts and prototype interactive visualizations. IEEA extends existing EEA frameworks with new analytic and interactive techniques that fundamentally enable new capabilities and insights to be derived from EEA, resulting in superior dynamic strategies for sustainment of system value delivery. These extensions enable the framing and analysis of large-scale problems, such as those posed by the DoD's Engineered Resilient Systems (ERS) efforts.

Future work will further extend interactive techniques to allow for improved analyses and decision-making. Improvements to IEEA processes for multi-era analysis, which has been covered less in existing literature than single and multi-epoch analysis, is an especially important area of future research. Active research in this area includes potential extensions of existing metrics, visualizations and analytical methods.

---

### **DISCUSSION**

The fundamental question at the focus of this research is, "How do we design, implement and operate systems that can sustain value delivery when there is uncertainty in future stakeholder needs, operating context or system variation?" The potential for loss of value due to future uncertainty can be overlooked by traditional system design techniques such as tradespace exploration and multi-disciplinary optimization (MDO) that often assume that mission requirements and needs are stable over time. EEA, which was developed to model these types of lifecycle uncertainties, can also encounter some limitations when we try to scale it up to the

larger scale problems now posed by the DoD. Specifically, it can potentially result in large and complex data sets that are difficult to generate, visualize and perform analysis on. It is hypothesized that by incorporating recent research in visual analytics with EEA, superior methods for early-stage system concept selection will be enabled.

IEEA provides a useful approach for identifying systems that sustain value while controlling costs by bringing knowledge about future uncertainty in needs, context and system variation with time forward in the system lifecycle. Since the early conceptual design stage is often the phase of the system lifecycle when most funding is committed and there is typically the least knowledge this is a point of high leverage. The IEEA framework extends EEA constructs to demonstrate how systems with sustained value can be designed via an interactive application. The contributions of this research are important not only to the broader field of systems engineering research, but also to industry practitioners that are actively seeking methods to address uncertainty when designing complex systems.

---

#### **NEXT STEPS**

The researchers will continue to mature the approach for evaluating systems under dynamic uncertainty, with further development of the extended framework for interactive capability and scaling to big data. New research includes enhancing interactive capability, investigating additional visual analytic techniques and tools, and applying the maturing prototype framework with supporting tools in additional case analysis and application. An impact assessment for use of IEEA methods will be developed.

---

## MODEL CHOICE AND TRADEOFF

One of the key challenges identified in preliminary research for IMCSE involves understanding the role that model choice plays in the generation and analysis of data for decision making. IMCSE anticipates making a key contribution in terms of framing this challenge and insights gained when actively trading models as a part of a study.

---

## BACKGROUND

As evidenced by the recent rise of model-influenced systems engineering efforts, including Model-Based Systems Engineering, Model-Based Engineering, and Interactive Model-Centric Systems Engineering, the role of models in engineering activities has been increasing in scope<sup>135,136</sup>. Models have always been used as tools to augment human ability to make predictions or sense of information, encapsulating existing knowledge, as well as automating its application. The rapid rise of low expense computational ability has increased the accessibility of numerical models and the roles they can play in engineering, including both analysis and synthesis. Leveraging models in an effective way for engineering decision support necessitates understanding the role that model choice plays in the generation and analysis of data for decision making. This is especially true when seeking to identify system solutions in early design that are robust across uncertainties<sup>137</sup>. This report section describes research in helping to frame this challenge and potential insights that might be gained when choosing and actively trading models as a part of a study.

### Enabling Model Tradeoff through Choice

Models can be used to support decision making throughout a system's lifecycle. These models, however, must appropriately suit their purpose, especially when used early in the lifecycle when data may be scarce and ambiguity and uncertainty are large. In such situations, there is not only a question of "fidelity" and whether a model is suited for purpose, but also a human aspect to acceptance of such models. This human aspect relates to the tension between a model's trustfulness (i.e. "to what degree can a person believe the model?") and its truthfulness (i.e. "to what degree does the model accurately represent the reality it purports to predict?")<sup>138</sup>. Making sure a model meets these standards is a challenge and one that can be addressed through the concept of model tradeoff through choice.

There are several key concepts involved during design decision making in early phase design. Figure 19, presented in the Phase 2 report, depicts the general relationship between decision problems and decision solutions as they relate to data and models in early phase engineering

---

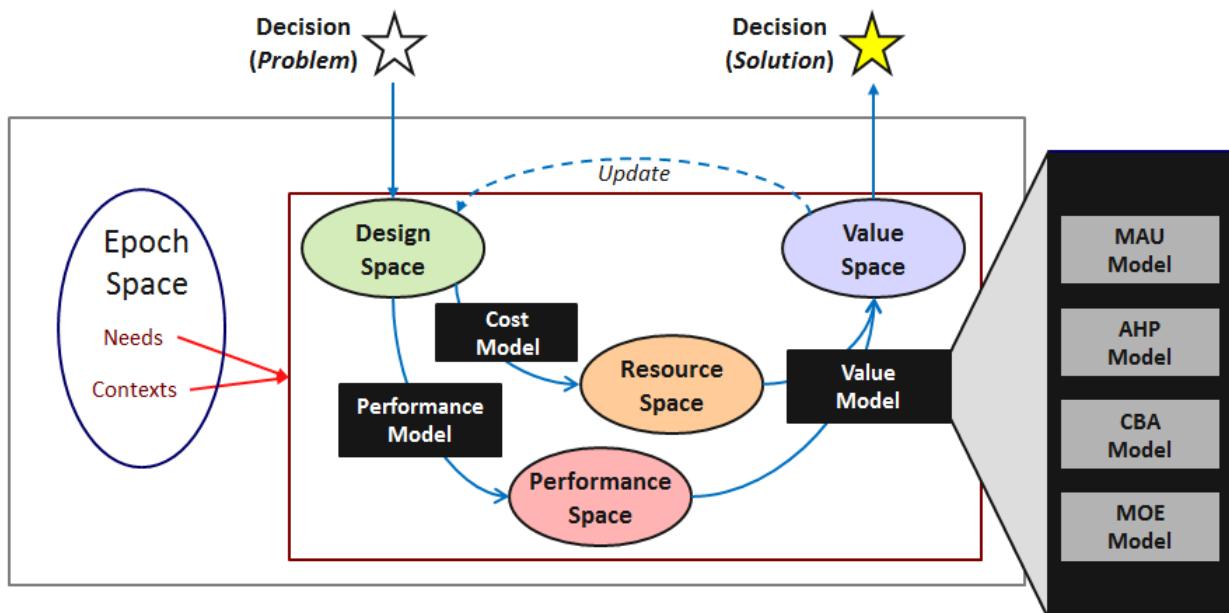
<sup>135</sup> Rhodes DH, Ross AM. Interactive Model-Centric Systems Engineering (IMCSE) Phase 1 Technical Report. SERC-2014-TR-048-1; September 2014.

<sup>136</sup> Rhodes DH, Ross AM. Interactive Model-Centric Systems Engineering (IMCSE) Phase 2 Technical Report. SERC-2014-TR-122-2; February 2015.

<sup>137</sup> Spero E, Avera MP, Valdez PE, Goerger SR. Tradespace exploration for the engineering of resilient systems. *12<sup>th</sup> Conf on Sys Eng Research. (CSER14)*. Redondo Beach, CA, March 2014.

<sup>138</sup> Ricci, N., Schaffner, M.A., Ross, A.M., Rhodes, D.H., and Fitzgerald, M.E., "Exploring Stakeholder Value Models Via Interactive Visualization," 12th Conference on Systems Engineering Research, Redondo Beach, CA, March 2014.

analysis. In this figure, decision problems suggest a space of potential solutions, which span a design space. The design space is then sampled and evaluated through two types of models: cost models and performance models. Cost models seek to predict the resources needed to develop and operate each of the evaluated potential systems. Typically, these estimates are in terms of dollars, and potentially time (i.e. schedule). Performance models seek to predict the operational behavior in context of the evaluated potential systems. Collectively we can term the cost and performance models as “evaluative” models, as they seek to evaluate (i.e., assign metrics to) alternatives. Value models seek to map the resulting resource and performance predictions into decision-friendly perceived benefit and cost metrics. Value models can be simple (e.g., just the cost and performance measures), or complex (e.g. aggregate perceived benefit and cost under uncertainty of a large number of measures), with many possible implementations<sup>139</sup>. Each of these models, and the artificial data generated by them, can be potentially altered by changes in the epoch space (i.e., exogenous context and needs changes). Updating occurs when users seek to modify the space definitions, or the models, in order for them to better address the problem under consideration (or to improve the trust or truthfulness (i.e., validity) of the models and data).



**Figure 19. Role of key models for supporting system decision making, with alternative value models use in demonstration case**

Since the role of models is central in the depicted decision framework, it is essential that engineers and analysts understand not only the sensitivities of their proposed solutions, but also of the models from which the data for decisions are generated. This includes understanding the impacts of key assumptions and model formulations on the data. One means for conducting this investigation is through “model trading” (i.e., model selection) where data is generated using

<sup>139</sup> Lee BD, Binder WR, Paredis CJJ, “A Systematic Method For Specifying Effective Value Models,” CSER14. March 2014.

alternative models with the resulting data compared. Another key consideration is understanding how the decisions are framed, both within the larger context of the engineering endeavor and within the specific context of the decision at hand.

---

#### **DEMONSTRATION OF EVALUATIVE MODEL TRADING: SPACE TUG**

In the Phase 2 report, the research team presented a demonstration of value model trading using the Space Tug case. For this exploratory case, we focus on evaluative models. The problem is framed as the following:

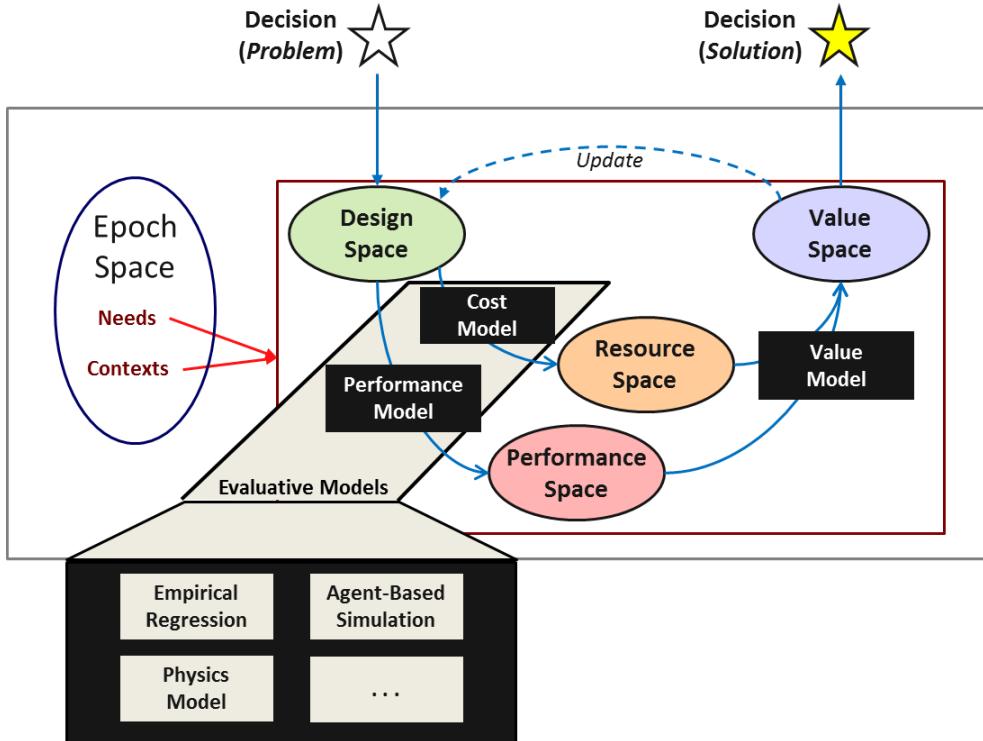
Models have increased in importance for engineering practitioners with the continued rise of powerful and accessible computation and, increasingly over the past few years, researchers have sought to develop standards and procedures allowing models to be leveraged more effectively<sup>140</sup>. Models are used to rapidly and accurately assess potential system concepts early in the design lifecycle and to explore the value tradeoffs associated with decisions designers can choose to make. Additionally, models can perform tasks that humans may struggle with, including the consideration of hypothetical scenarios that form the basis of uncertainty<sup>141</sup>. This demonstration discusses continuing research results in the area of model comparison and trading: the active comparison of different model-generated results in order to learn more about the system of interest.

Figure 20 illustrates an alternative to the earlier conceptual framework for the decision process in early system design, including the relationships between the various models that are used to support the decision maker. The general flow involves the creation of a design space suited to the problem in which each design is evaluated using a set of evaluative models to determine its performance and cost (i.e., resources required) with respect to a set of given contextual factors. Those performance and resource attributes are then fed into a value model in order to assess the “goodness” of each alternative, which is the key decision-making criterion.

---

<sup>140</sup> Rhodes DH, Ross AM. Interactive Model-Centric Systems Engineering (IMCSE) Phase 1 Technical Report. SERC-2014-TR-048-1; September 2014.

<sup>141</sup> Spero E, Avera MP, Valdez PE, Goerger SR. Tradespace exploration for the engineering of resilient systems. 12th Conf on Sys Eng Research. (CSER14). Redondo Beach, CA, March 2014.



**Figure 20. Role of key models in system decision making, with example evaluative models called out.**

The role and impact of models on the design process is a core interest of the Interactive Model-Centric Systems Engineering community. In our previous demonstration, we expanded the earlier concept of interactively refining a value model<sup>142</sup> into the potential use of value model trading to support the decision making process<sup>143</sup>. Specifically, the ability to compare the tradespace as it exists under multiple value models was shown to be a powerful means for building trust in the model results, particularly when a decision maker may be unsure how to mathematically represent their needs early in the system lifecycle. This demonstration addresses a similar model trading concept, this time for the evaluative models.

Evaluative models come in a plethora of different forms, too many to list exhaustively. A few common examples of evaluative models are provided in Figure 20. Depending on the system in question, different types of evaluative models will be appropriate and/or available. Commonly discussed characteristics of evaluative models include fidelity (or the similar concepts of accuracy and precision) and computational costs (among others such as purpose and credibility<sup>144</sup>). When choosing models based on fidelity and computational cost, there is often a tradeoff, so the “best” model for a given task may be subjective. This choice may also require the consideration of the confidence the decision maker has in each model, which may not be solely determined by its fidelity. This best-model-calculus and associated tradeoffs are interesting topics and are likely

<sup>142</sup> Ricci N, Schaffner MA, Ross AM, Rhodes DH, Fitzgerald ME. Exploring stakeholder value models via interactive visualization. 12th Conf on Sys Eng Research. (CSER14). Redondo Beach, CA, March 2014.

<sup>143</sup> Ross AM, Rhodes DH, and Fitzgerald ME, "Interactive Value Model Trading for Resilient Systems Decisions," 13th Conf on Sys Eng Research. (CSER15). Hoboken, NJ, March 2015.

<sup>144</sup> NASA. Standard for Models and Simulations. NASA-STD-7009; July 2008.

what most readers will think of first when hearing the phrase “evaluative model trading” (e.g., determining a model’s “fit for purpose” and associated verification, validation, and accreditation (VV&A) activities<sup>145</sup>). However, this demonstration is not concerned with *selecting* the best evaluative model from a set of choices, but rather *leveraging* the ability to use multiple evaluative models in order to support the decision process. That is, instead of expending efforts to find the “right” model, what can be determined by leveraging multiple different models in order to garner potentially novel insights, especially when “fit for purpose” may be unclear early in the system lifecycle?

Why might engineers be interested in using multiple evaluative models? On a basic level, running multiple evaluative models and comparing their results can support cross-validation of each model and increase decision maker confidence in their results. Of particular interest to this research is the use of models to support early concept decision making, which may require measuring the expected performance of new or emerging technologies that have yet to be built or tested. In this case, it is possible that no evaluative models are truly validated, leading to a situation similar to that of the value model trading problem, but instead of having no “ground truth” to validate against (since value is subjective), the designers simply *don’t know* what that ground truth is. As a result, searching for alternatives that are robust to the unknown accuracy or precision of the models is a powerful use of multiple models. The following case demonstrates this concept on a simple example.

### **Space Tug Demonstration overview**

In order to demonstrate the effects of trading evaluative models and the insights that can be gained by doing so, we will return to the Space Tug case used in the prior demonstration of value model trading. The generic mission is therefore the same but the key questions have changed:

A decision maker has a budget for an orbital transfer vehicle (a.k.a. “Space Tug”) and knows what he wants (in terms of attributes of goodness of a system). However, he is aware that Space Tugs are a developing technology, and the models used to evaluate them are not 100% validated. He therefore wants to explore a variety of model implementations in order to understand the following:

1. How do changes in the model impact the apparent “best” solutions?
2. Are there system designs that are robust to changes in the model?
3. What patterns in the performance space are driven by model artifacts?

In addition to these questions, a decision maker may also be interested in exploring the tradeoffs between model fidelity, confidence in decisions, and computational effort. However, in this simple example, the models have effectively zero computational cost and therefore we will focus on the implications to the system design only.

---

<sup>145</sup> INCOSE. Systems Engineering Handbook: A Guide for system life cycle processes and activities. 4th ed. Chapter 9.1 Modeling and Simulation. INCOSE-TP-2003-002-04. 2015.

## Models used in the case

The value model is held constant in this demonstration, using the multi-attribute utility model described in the previous study. Four different evaluative model implementations were tested and compared. These implementations were not created to teach specific lessons; all insights gained from their analysis were emergent.

### Implementation #1 (original model)

The Space Tug's default evaluative implementation is a combination of models<sup>146</sup> including a manipulator capability lookup table, a binary fast/slow speed assignment, linear models for mass and cost, and the rocket equation. There are four design variables (input elements under designer's control) leading to 384 different design alternatives, and four evaluated attributes (model outputs corresponding to the value of the system). Figure 21 shows the default model implementation and resulting benefit-cost tradespace, with the Pareto set marked with blue triangles. Only 83 of the 384 evaluated alternatives are feasible, for a yield of approximately 21.6%. This is identical to the multi-attribute utility tradespace in the value model trading study.

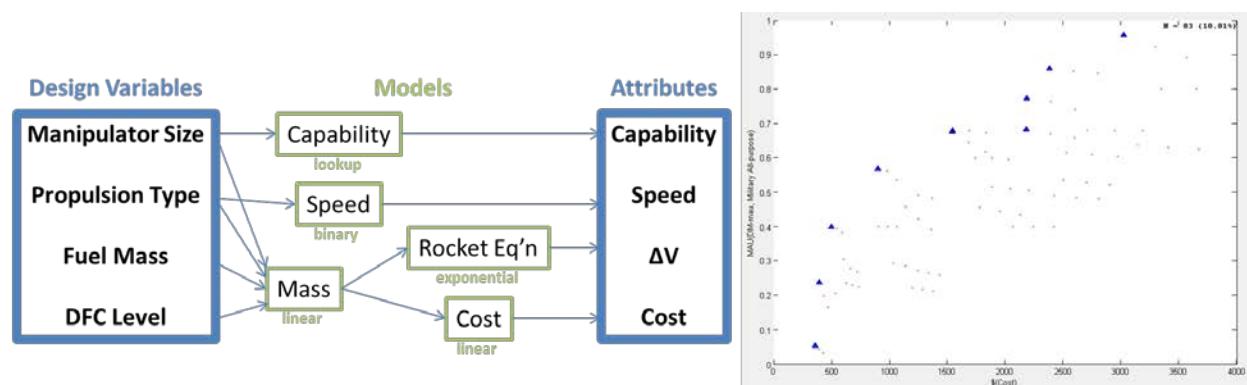


Figure 21. Space Tug model implementation #1 and resulting tradespace.

### Implementation #2 (new speed model)

The second implementation is a classic fidelity upgrade. The binary speed model has been replaced by an acceleration model based on the thrust of the chosen propulsion type and the mass of the system using the classic  $F=m*a$  formulation. This model provides a more accurate estimate for the resulting speed of the system. Note that this change does not impact the yield: this is because no alternatives were designated infeasible using the original speed model but feasible using the new model (or vice-versa). Figure 22 shows this implementation and tradespace, with the Pareto set marked in green, left-pointing triangles.

---

<sup>146</sup> McManus H. Schuman T. Understanding the orbital transfer vehicle trade space. AIAA Space 2003. Long Beach, CA, September 2003.

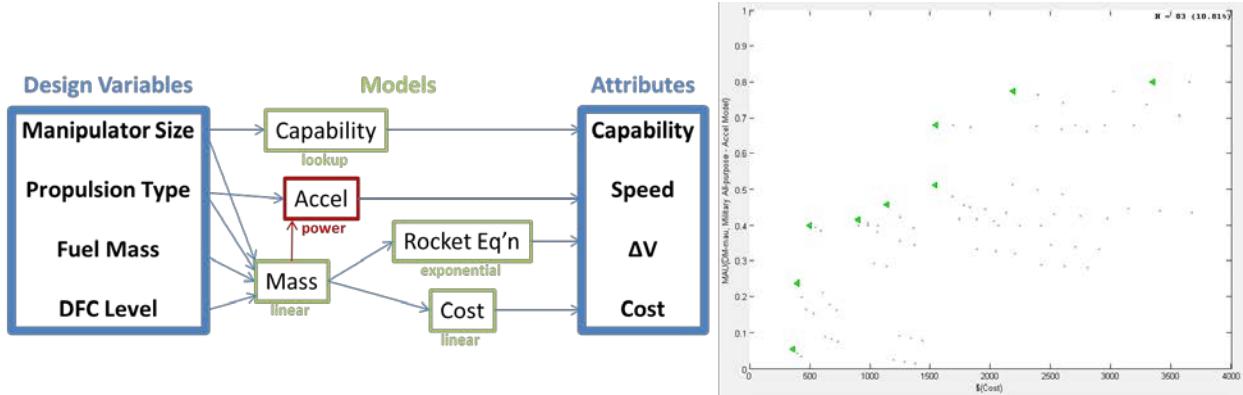


Figure 22. Space Tug model implementation #2 and resulting tradespace.

### Implementation #3 (new material model)

The third implementation is a “representational” fidelity upgrade. The mechanics of the models remain fixed but the tradespace has been expanded by taking the material comprising the support structure, previously assumed to be aluminum, and making it a design variable called *material*, accepting “aluminum” or “carbon” as choices. This allows for a more detailed representation of the system through the inclusion of other material options into the tradespace. Now the tradespace doubles in size to 768 alternatives by adding the choice of a carbon structure. The mass and cost models must accommodate this change, as carbon has a lower density but higher cost than aluminum. With this model implementation the yield *does* change, increasing to approximately 23% thanks to designs that, while infeasible with aluminum structure due to insufficient delta-V, are feasible with the lighter carbon structure. Figure 23 shows this implementation and tradespace, with the Pareto set marked in cyan, down-pointing triangles.

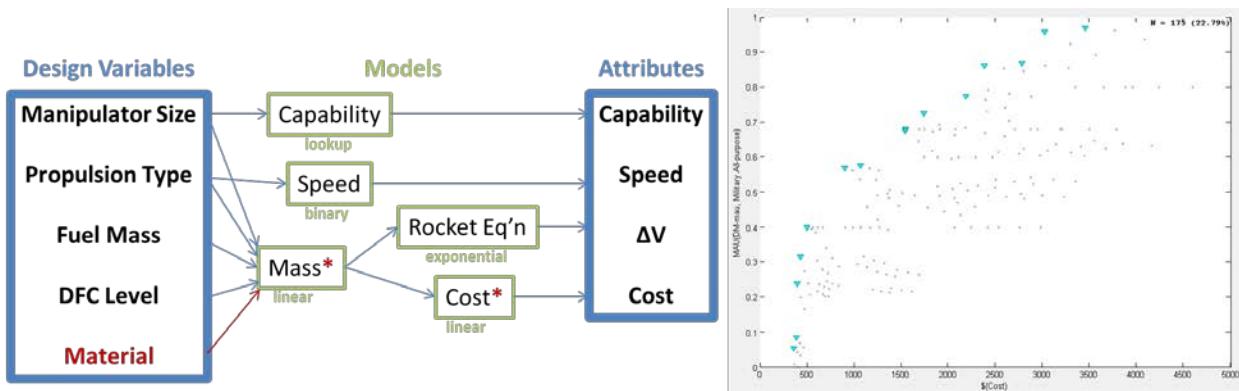


Figure 23. Space Tug model implementation #3 and resulting tradespace.

### Implementation #4 (new speed and material model)

The final implementation combines the model changes in #2 and #3, having a more detailed acceleration model *and* the expanded tradespace. Evaluative models can be hybrid in nature, allowing for mixing and matching sub-models as desired. Figure 24 shows this implementation and tradespace, with the Pareto set marked in red, right-pointing triangles.

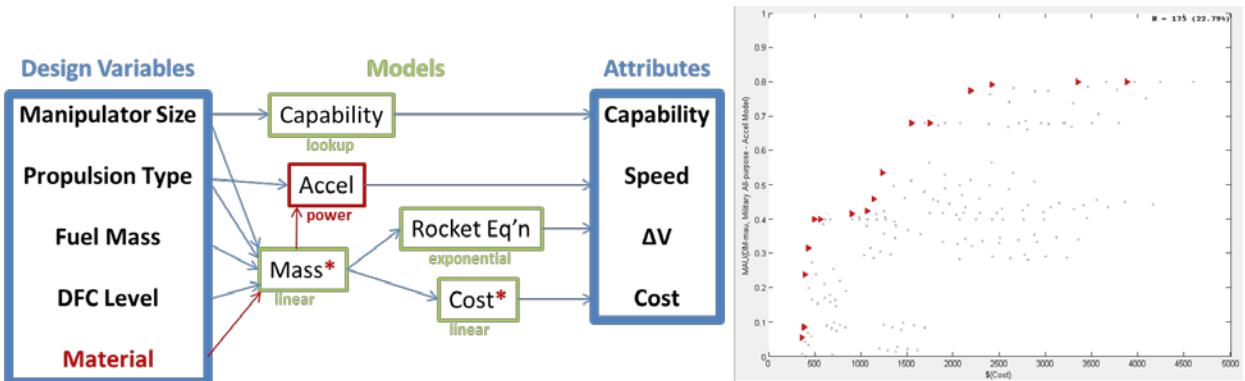


Figure 24. Space Tug model implementation #4 and resulting tradespace.

## Results

### Comparisons via Pareto sets

Figure 25 shows all four tradespaces with all four Pareto sets marked. On inspection, there appears to be more overlap between the sets for this evaluative model trading study than for the prior value model trading study. This is a positive result: while value models can drastically reorder the value of different alternatives due to differing subjective interpretations, it would be indicative of poor “fit for purpose” if the evaluative models presented dramatically different estimates of the key performance attributes. In particular, there is considerable agreement in the Pareto sets in the low-cost domain of the tradespace, suggesting that the differences between these implementations are experienced mostly in larger, more expensive systems. Fuzzy Pareto set analysis was conducted as it was in the prior demonstration, but is omitted here for two reasons: (1) designs that are jointly efficient between implementations already exist at zero fuzziness and (2) the resulting patterns between implementations are extremely similar to those identifiable in the basic Pareto sets, but more difficult to “see” when plotted on the tradespace. For the record, considering a 1% fuzzy Pareto set approximately doubles the number of designs under consideration in each model implementation, while 5% approximately triples it.

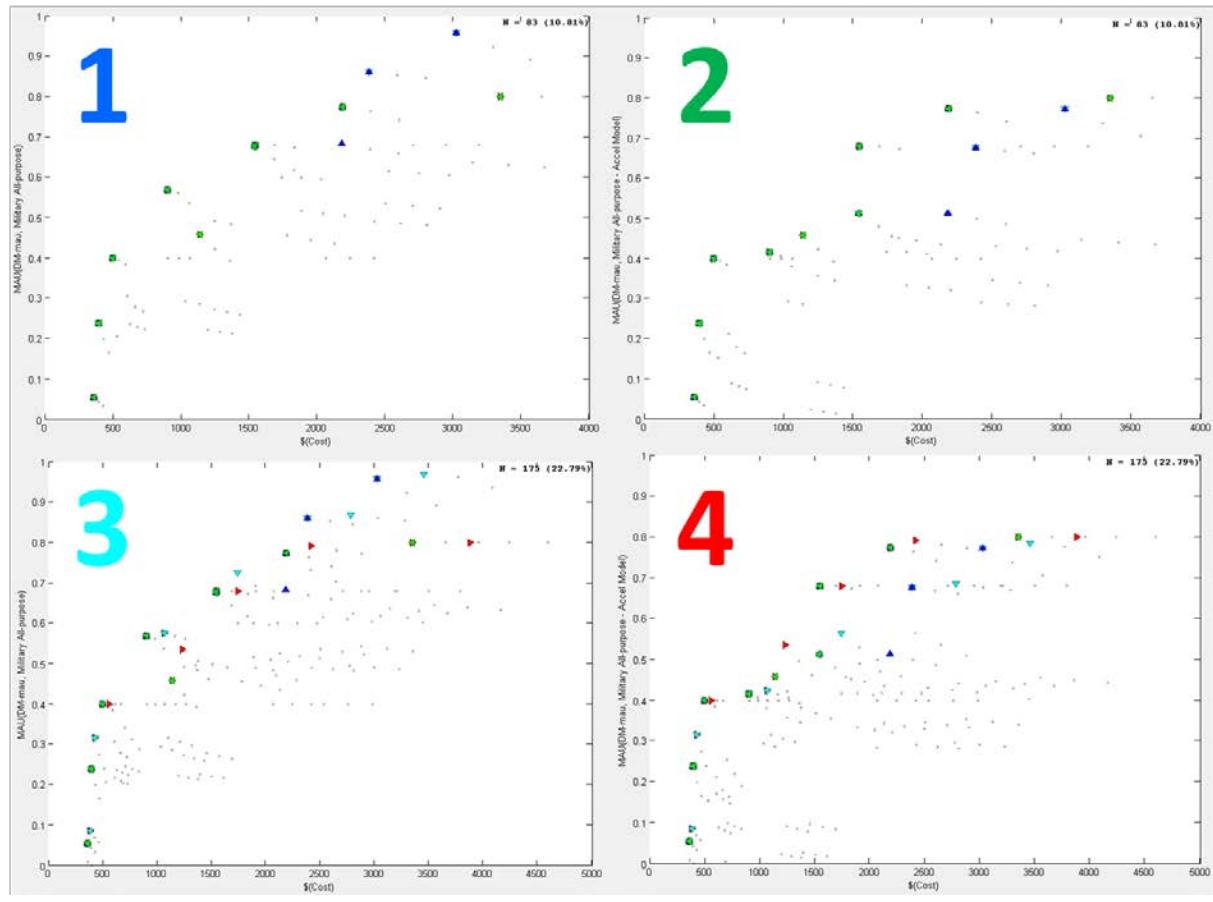


Figure 25. The tradespaces for all four implementations, with all four Pareto sets marked in each

### Joint Pareto analysis

To explore these designs in more detail, Table 12 includes all of the Pareto-efficient marked designs in the above figures and shows in which implementations they are efficient. Many designs that are efficient with aluminum material are also efficient when changing to carbon in model implementations 3 and 4, therefore the table is set up such that designs that are identical except for material are placed in the same row. From this list there are six emergent categories of efficient designs, which we will identify as A through F:

- **A.** Designs 52, 53, and 63 (and their carbon counterparts) are always efficient, in every implementation. These alternatives are robust to evaluative model trading, and are all among the low-cost solutions previously identified by inspection of the scatterplots.
- **B.** Designs 54, 87, and 119 are always efficient except for their carbon variants under the binary speed model in implementation 3. All of these alternatives have electric propulsion (categorized as “slow” by the binary model) and delta-V above or very near the maximum-utility point in the utility function (20,000 m/s). This results in no additional benefit when switching to carbon, but the additional costs are still

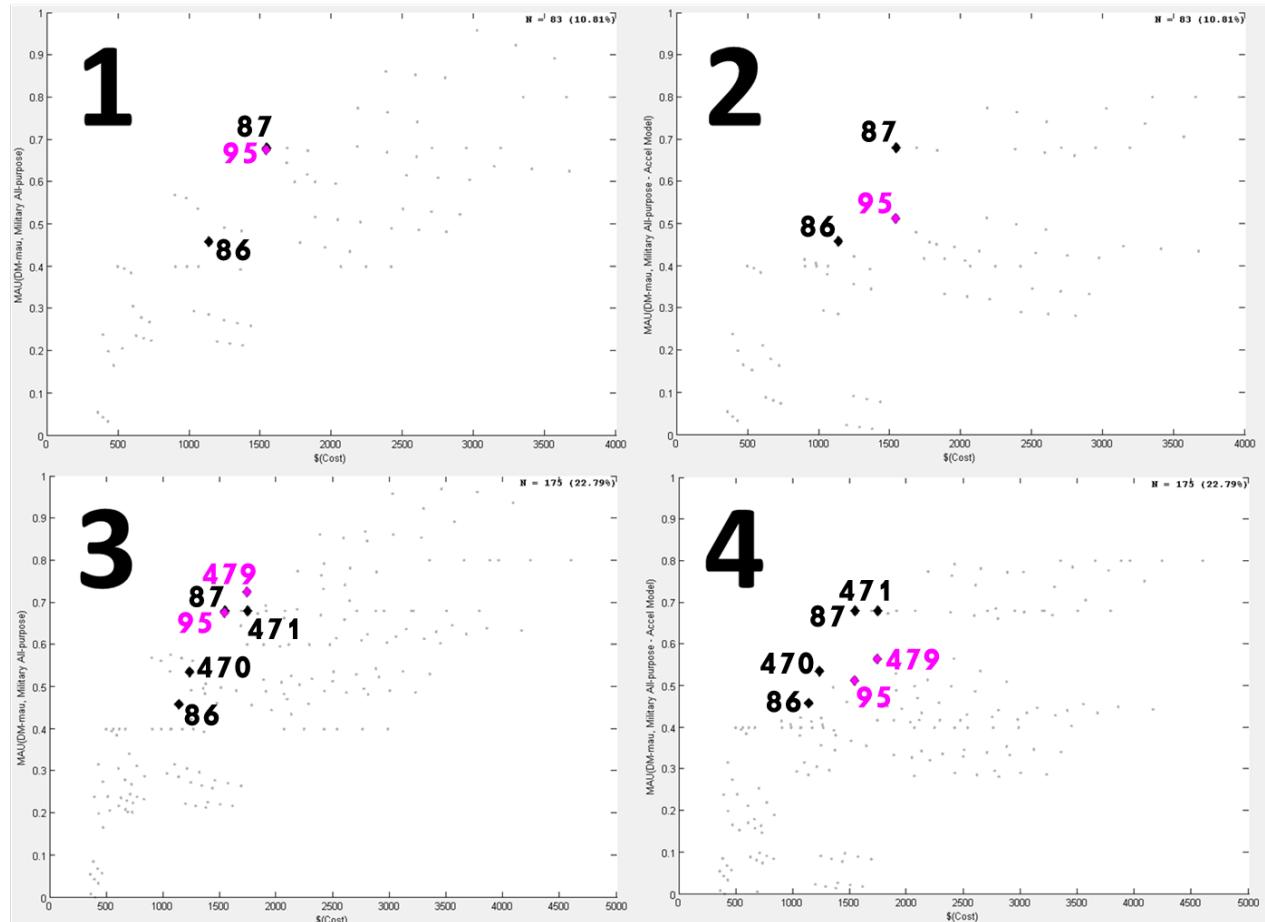
experienced. In implementation 4 however, the reduced weight from the carbon structure leads to improved acceleration, making these designs efficient again.

**Table 12. Designs, marked in gray and with a check for model implementations in which they are efficient**

Category	Design ID (Aluminum)	Model 1	Model 2	Model 3	Model 4	Design ID (Carbon)	Model 3	Model 4
A	52	✓	✓	✓	✓	436	✓	✓
	53	✓	✓	✓	✓	437	✓	✓
	63	✓	✓	✓	✓	447	✓	✓
B	54	✓	✓	✓	✓	438		✓
	87	✓	✓	✓	✓	471		✓
	119	✓	✓	✓	✓	503		✓
C	86		✓		✓	470		✓
	120		✓		✓	504		✓
D	96	✓		✓		480	✓	
	128	✓		✓		512	✓	
E	127	✓				511		
F	95	✓	✓	✓		479	✓	

- C.** Designs 86 and 120 are efficient in both materials but only under the improved speed model of implementations 2 and 4. These designs are the same as 87 and 119 except for slightly more and slightly less fuel, respectively. The higher fidelity acceleration calculation results in slightly different “sweetspots” in the tradeoff between speed and delta-V driven by fuel mass.
- D.** On the other hand, designs 96 and 128 are efficient in both materials but only under the original speed model in implementations 1 and 3. These designs both have the maximum fuel mass enumerated in the tradespace. This extra mass does not penalize speed under the original model, allowing these designs to be efficient; however, the switch to the acceleration model lowers their value and removes them from the Pareto front.
- E.** The above categories have physically-intuitive explanations for their behavior, but some insights from model trading do not. Trading models can also show that some conclusions that could be drawn from a single tradespace are nearly entirely artifacts of the model implementation. For example, design 127 is efficient only in the original implementation. This design is the same as 128 but with less fuel, and sits in a noticeably concave region of the Pareto front, which is surpassed by many designs when considering higher fidelity and/or alternative material models. Model trading reveals this design to be much less attractive than it originally appears.
- F.** Perhaps the most unintuitive insight of all is that design 95 and its carbon counterpart 479 are efficient in implementations 1, 2, and 3 but *not* 4. It is not immediately apparent why a given system design would be efficient using the original models and higher fidelity models, but not when those higher fidelity models are combined,

especially for a simple case such as Space Tug where most of the equations in play are linear. Figure 26 shows the four tradespaces with 95/479 highlighted in magenta and 86/87/470/471 highlighted in black. In implementation 1 (top-left), 95 is closely co-located with 87. The shift to implementation 2 and the acceleration model (top-right) causes 95 to decrease in benefit and occupy a concave region of the Pareto front between 86 and 87. On the other hand, shifting to implementation 3 (bottom-left) nearly replicates the implementation 1 pattern with the three carbon designs, shifted to slightly more cost and benefit but such that they still overlap the original three designs. However, combining these two effects in implementation 4 (bottom-right) results in overlapping concave patterns, with the “elbow” points in each triplet, designs 95/479, now dominated by an end point of the other triplet: 95 by 470 and 479 by 87. This illustrates the benefit of model trading for capturing deep insight into the tradespace, as this interaction indicates that, despite the inherent simplicity of this example, there is some unpredictable interplay between the speed and mass submodels that may merit further detailed analysis of their component functions.



**Figure 26. The tradespaces for all four implementations, highlighting the model artifacts impacting designs 95/479 (magenta), combining to make them inefficient in implementation 4 only. Design numbers left/above of the points are on the Pareto front, right/below are inefficient.**

The four “promising” designs identified by the value-model trading study (63, 95, 127, and 128) are all present in the above list, a fact that was guaranteed since the prior analysis confirmed their efficiency in the equivalent of implementation 1. However, it is interesting to note that each of these designs falls into a different category of interest across the different evaluative implementations. Combining the insights of the two studies suggests that designs 127 and 128 are not as good as originally believed, but 63 and 95 (or their carbon variants, which were not in the value-trading study) are still potentially interesting selections.

Finally, we can attempt to build some intuition for the impact of model implementation in terms of the design variables of the efficient alternatives. Figure 27 shows implementations 1 and 4 with the different Pareto efficient designs colored by which category they belong to in Table 12. In each tradespace, regions of the Pareto front with consistent patterns are highlighted. First, note that no bipropellant or cryogenically propelled Space Tugs are efficient: this is the result of the multi-attribute utility function, which was noted in the prior study to eliminate most of those alternatives due to high requirements on delta-V. However, we can see here that the relative value of the remaining Pareto efficient electric and nuclear designs is sensitive to model implementation.

While both tradespaces have medium payload, electric and nuclear Space Tugs in the low cost region (members of category A and the low end of B), there are two main impacts of the model changes. First, in the medium cost region, implementation 4 has a convex region that strongly favors the electric designs of categories B and C, whereas implementation 1 is nearly linear and features crossover between the electric designs and large nuclear designs of F. Second, the extra-large nuclear designs of category D dominate the high cost region in implementation 1 but switch to high-end electric designs of C in implementation 4. Overall then, it seems that the new acceleration model favors electric designs more than the original model by closing the gap on both ends – giving electric designs more credit for acceleration than a binary ‘0’ and penalizing high-mass nuclear designs. This means that electric designs are more robust to these model trades, while the regions of the tradespace for which nuclear designs are optimal will change more dramatically depending on implementation. However, it is worth pointing out again that reordering of these designs is less dramatic than it was for value model trading, as this type of analysis captures slight differences in efficient and nearly-efficient designs, rather than a complete redefinition of value.

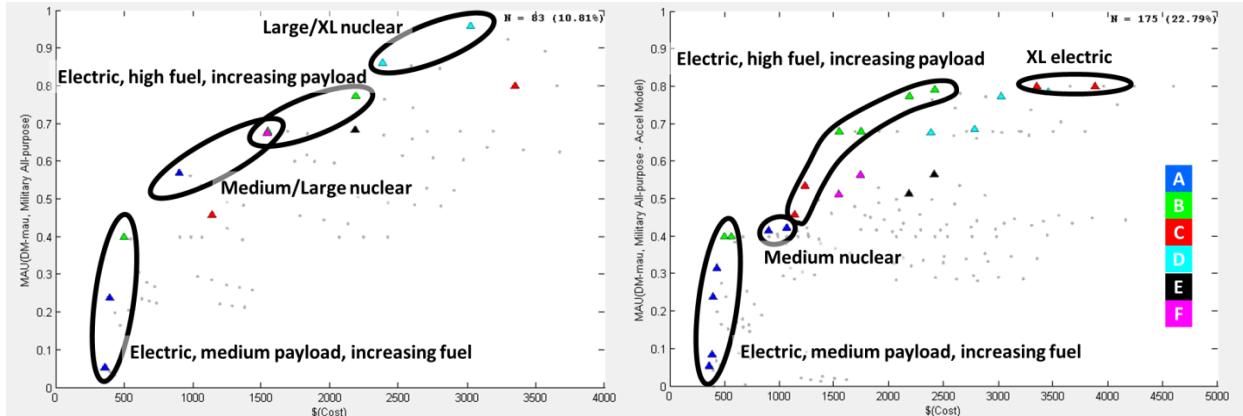


Figure 27. The tradespaces for implementations 1 (left) and 4 (right), highlighting design variable patterns on the Pareto front. The six categories of designs in Table 12 are used as the color scheme.

### Discussion of Evaluative Model Case

Designers can gather insight about their preferences from the exploration of value model trades, which can be a useful exercise given the lack of ground-truth data to support their validation. Evaluative performance and cost models can similarly be traded and explored, which may become particularly worthwhile for early-concept designs involving emerging technology too new to have established and validated supporting models. These evaluative model trades often take the form of fidelity differences in either representation (of the design) or evaluation (of the system attributes), and typically come with an associated difference in computational cost. In some cases, model trades may encompass completely different physical phenomenologies, such as when engineers must select a turbulence model within a larger fluid dynamics model – a decision that is often made with the support of experimental prototyping data to identify the model that is most accurately predicting the resulting flow. In this way, evaluative model trading can support both the identification of approximate bounds or errors on the results of model-centric engineering efforts and build designer trust and confidence in data by revealing the impacts of switching model components and the idiosyncrasies that can arise at the intersection of multiple models. The demonstration case provided here was intended to focus on the latter of these points, as building of model trust is considered a strength of the tradespace exploration paradigm. Future studies can also explore the tradeoff between model fidelity, designer confidence, and computational effort, which will require a case with a more complex model with a higher available top-end fidelity and longer resulting computation times.

---

### PRELIMINARY DEMONSTRATION OF COMBINED VALUE MODEL AND EVALUATIVE MODEL TRADING: SPACE TUG

The combination of both value model trading and evaluative model trading is anticipated to provide further insights into the impact of model choice, and emergence within such models, on potentially attractive design solutions. This section now describes a preliminary demonstration of combined value and evaluative model trading. The next phase of research will further refine the combination the two into a coherent model trading framework that allows for engineers to explore the impacts of uncertainty in both domains at once. This research will also be integrated with other efforts of the developing Interactive Model-Centric Systems Engineering research

effort in support of its larger goals for effective, integrated modeling and exploration environments and practices<sup>147,148</sup>.

### Models used in Case

For the combined demonstration, we used both sets of value and evaluative models described in our prior demonstrations, listed in Table 13. Value models included multi-attribute utility (MAU), analytic hierarchy (AHP), cost-benefit analysis (CBA), and measure of effectiveness (MOE). Evaluative models included four implementations including the original (#1), new speed (#2), new material (#3), and combined new speed and material (#4). The details of these models can be found in their respective SERC report sections, or associated CSER papers.<sup>149,150</sup>

**Table 13 Value and evaluative models used in demonstration**

Evaluative Value	Implementation #1 (Original)	Implementation #2 (New Speed)	Implementation #3 (New Material)	Implementation #4 (New Speed and Material)
MAU	1-MAU	2-MAU	3-MAU	4-MAU
AHP	1-AHP	2-AHP	3-AHP	4-AHP
CBA	1-CBA	2-CBA	3-CBA	4-CBA
MOE	1-MOE	2-MOE	3-MOE	4-MOE

### Results

The results of this demonstration will be presented as a value model by evaluative model and evaluative model by value model comparison. That is, we repeated our value model tradeoff study for *each* of the evaluative models (4 value models x 1 evaluative model) x 4 evaluative models. Then we repeated our evaluative model tradeoff study for each of the value models (4 evaluative models x 1 value model) x 4 value models. We did this in order to demonstrate generalization of the approach we took within the prior studies, and to highlight the intent as one that aims to gain knowledge about the impact of model choice, rather than selection of the “best” alternative. Further work in IMCSE will refine the approach into a generalizable framework for conducting combined model tradeoff studies.

### Value Model Trading for each Evaluation Model

Recall from the earlier value model trading case that there are four pairs of objectives considered when determining Pareto efficient design sets. These are shown in Figure 28 below, with each value model resulting in a metric quantifying the expected benefit and cost of an alternative.

<sup>147</sup> Reymondet L, Rhodes DH, and Ross AM, “Considerations for Model Curation in Model-Centric Systems Engineering,” IEEE Syscon 2016, Orlando, FL, April 2016.

<sup>148</sup> Rhodes DH, and Ross AM, “A Vision for Human-Model Interaction in Model-Centric Systems Engineering,” INCOSE Int’l Symp 2016, Edinburgh, UK, July 2016.

<sup>149</sup> Ross, A.M., Rhodes, D.H., and Fitzgerald, M.E., "Interactive Value Model Trading for Resilient Systems Decisions," 13th Conference on Systems Engineering Research, Hoboken, NJ, March 2015.

<sup>150</sup> Ross, A.M., Fitzgerald, M.E., and Rhodes, D.H., "Interactive Evaluative Model Trading for Resilient Systems Decisions," 14th Conference on Systems Engineering Research, Huntsville, AL, March 2016.

Each value model has intentionally kept the cost metric as a distinct objective in order to explicitly highlight the cost versus benefit tradeoffs that determine value. This is reflected in the objective sets each having cost as well as the appropriate value model metric for benefit.

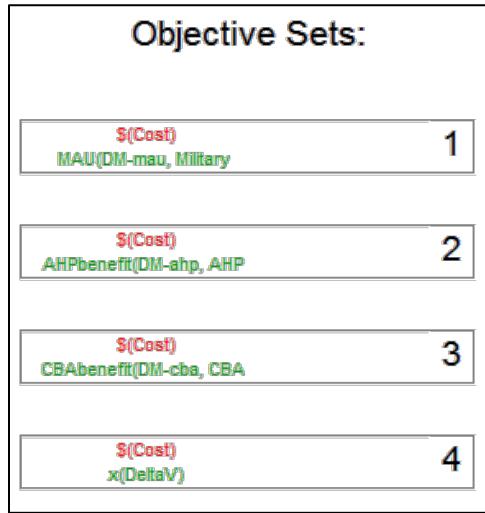


Figure 28. Four objective sets of two objectives each

#### Implementation #1 (original model)

Recall evaluative model #1 has a potential tradespace size of 384 designs, enumerated across three design variables (payload size, propulsion type, and fuel tank size). Table 14 describes the size of the 0% fuzzy Pareto sets for each of the value models when using evaluative model #1 (the original model). This exactly matches the original value model trading case, as to be expected.

Table 15 lists the designs that are almost joint Pareto efficient, appearing in 3 out of 4 Pareto sets.

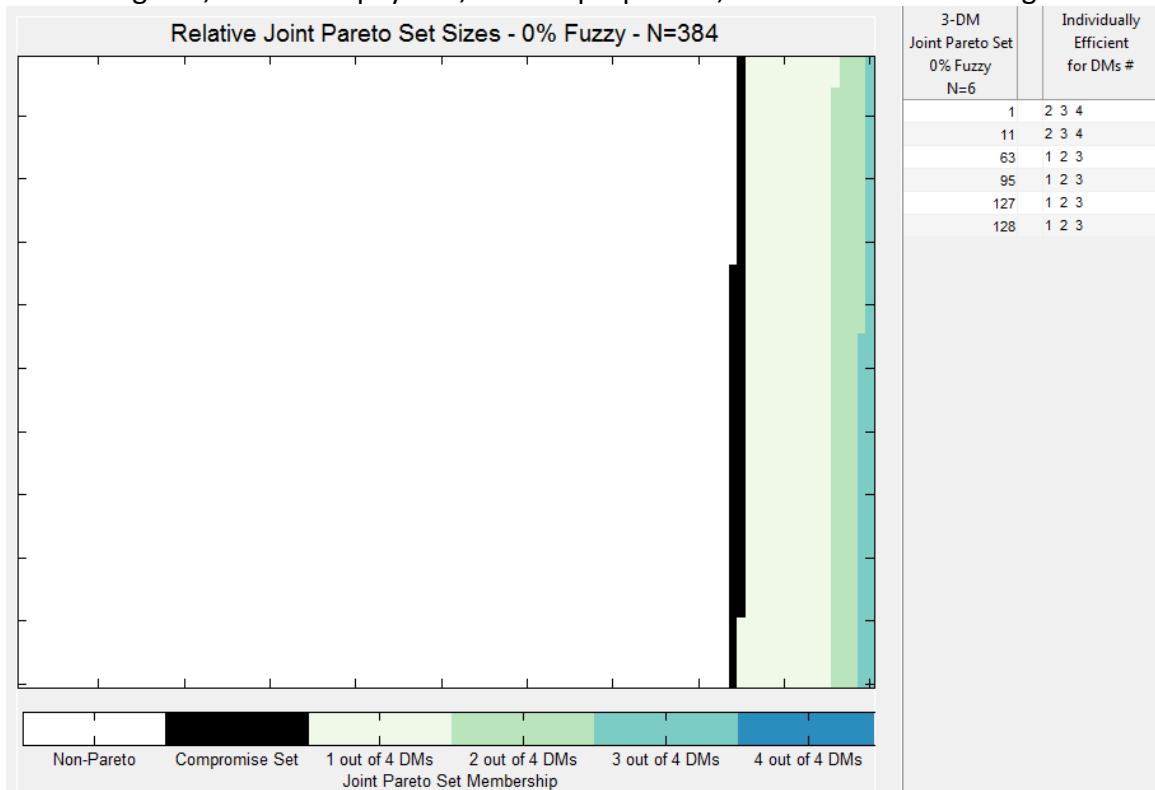
Table 14. Pareto sets sizes for value models using evaluative model #1

Objectives Set	# designs in 0% Pareto Set
1: MAU-COST	10
2: AHP-COST	43
3: CBA-COST	17
4: MOE-COST	13
JOINT	0
COMPROMISE	6

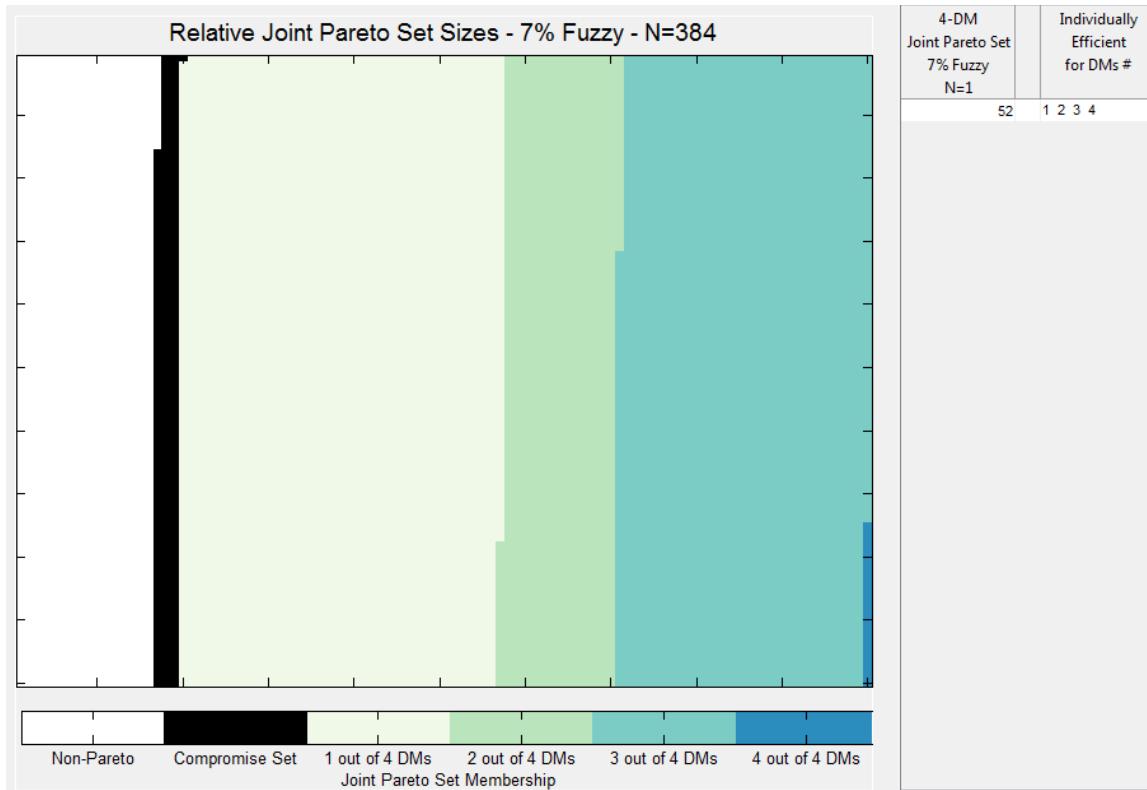
**Table 15. Promising designs that are joint Pareto efficient across 3 out of 4 value models in evaluative model #1**

ID Number	Pareto Efficient For	Invalid For	Details
1	2, 3, 4	1	small biprop min fuel
11	2, 3, 4	1	small cryo near min fuel
63	1, 2, 3		med nuke near max fuel
95	1, 2, 3		large nuke near max fuel
127	1, 2, 3		xl nuke near max fuel
128	1, 2, 3		xl nuke max fuel

First fully joint (across 4 out of 4 value models) Pareto efficient design appears at fuzzy level of 7% is design 52, a medium payload, electric propulsion, medium fuel tank design.



**Figure 29. Gridmap showing relative sizes of 0% fuzzy joint Pareto sets in evaluative model #1**



**Figure 30.** Gridmap showing relative sizes of 7% fuzzy joint Pareto sets in evaluative model #1

Figure 29 and Figure 30 illustrate the relative sizes of the fuzzy joint Pareto sets at 0% and 7% fuzzy levels respectively.

### Implementation #2 (new speed model)

Recall evaluative model #2 has a potential tradespace size of 384 designs, enumerated across three design variables (payload size, propulsion type, and fuel tank size). This model implementation has a new (higher fidelity) speed model. Table 16 describes the size of the 0% fuzzy Pareto sets for each of the value models when using evaluative model #2 (the new speed model). Table 17 lists the designs that are almost joint Pareto efficient, appearing in 3 out of 4 Pareto sets.

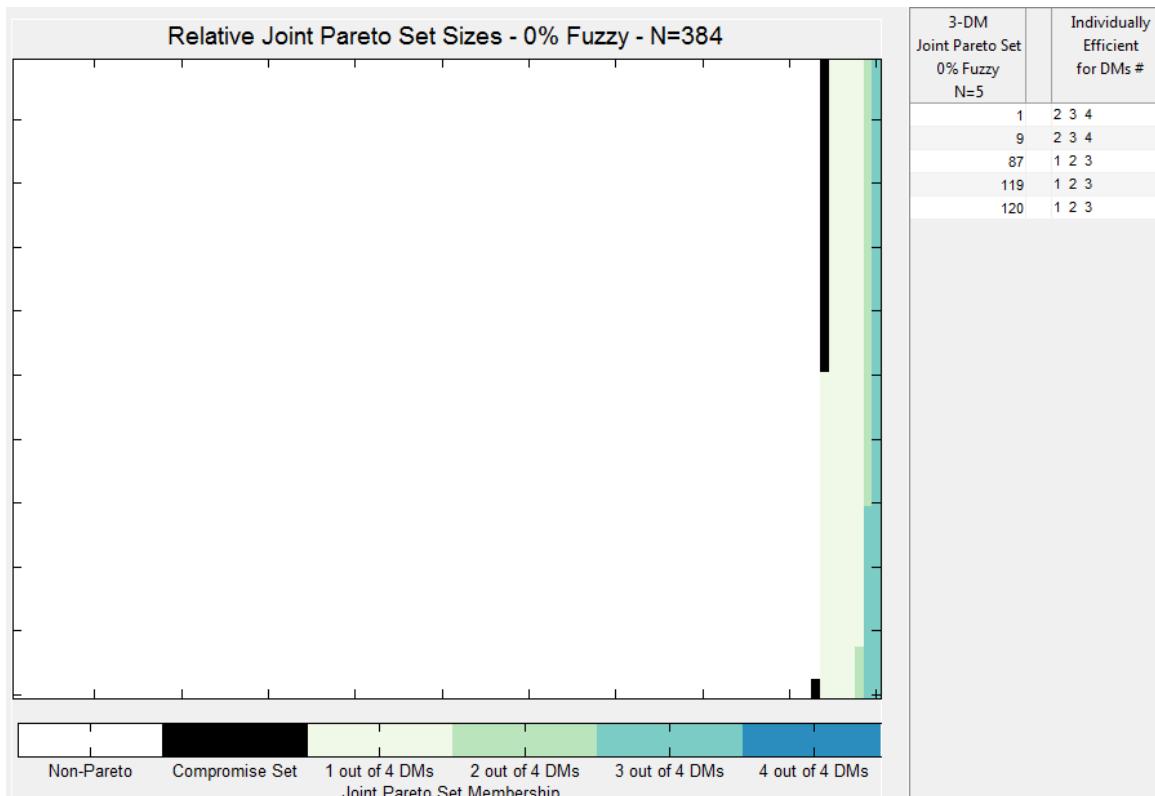
**Table 16.** Pareto sets sizes for value models using evaluative model #2

Objectives Set	# designs in 0% Pareto Set
1: MAU-COST	9
2: AHP-COST	10
3: CBA-COST	6
4: MOE-COST	13
JOINT	0
COMPROMISE	2

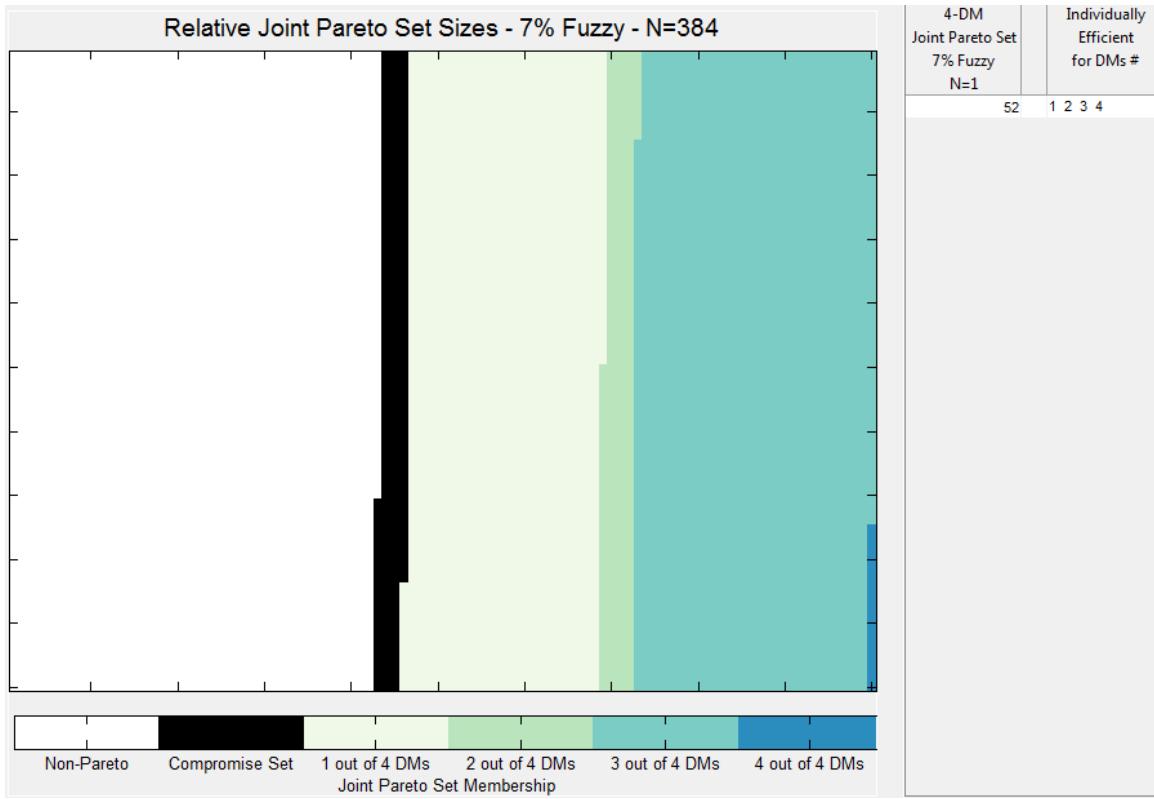
**Table 17. Promising designs that are joint Pareto efficient across 3 out of 4 value models in evaluative model #2**

ID Number	Pareto Efficient For	Invalid For	Details
1	2, 3, 4	1	small biprop min fuel
9	2, 3, 4	1	small cryo min fuel
87	1, 2, 3		large elec near max fuel
119	1, 2, 3		xl elec near max fuel
120	1, 2, 3		xl elec max fuel

The same design 52 appears as the first fully joint Pareto efficient at a 7% fuzzy level.



**Figure 31. Gridmap showing relative sizes of 0% fuzzy joint Pareto sets in evaluative model #2**



**Figure 32. Gridmap showing relative sizes of 7% fuzzy joint Pareto sets in evaluative model #2**

Figure 31 and Figure 32 illustrate the relative sizes of the fuzzy joint Pareto sets at 0% and 7% fuzzy levels respectively.

For this evaluative model it looks like the electric propulsion type passes the nuclear type in terms of “promising” for the new speed model. The first joint Pareto design appears at the same fuzzy level as the original evaluative model.

### Implementation #3 (new material model)

Recall evaluative model #3 has a potential tradespace size of 768 designs, enumerated across four design variables (payload size, propulsion type, fuel tank size, and material type). This model implementation added “fidelity” to the design space description, allowing for variation in the material type between aluminum and carbon.

Table 18 describes the size of the 0% fuzzy Pareto sets for each of the value models when using evaluative model #3 (the new material model). Table 19 lists the designs that are almost joint Pareto efficient, appearing in 3 out of 4 Pareto sets.

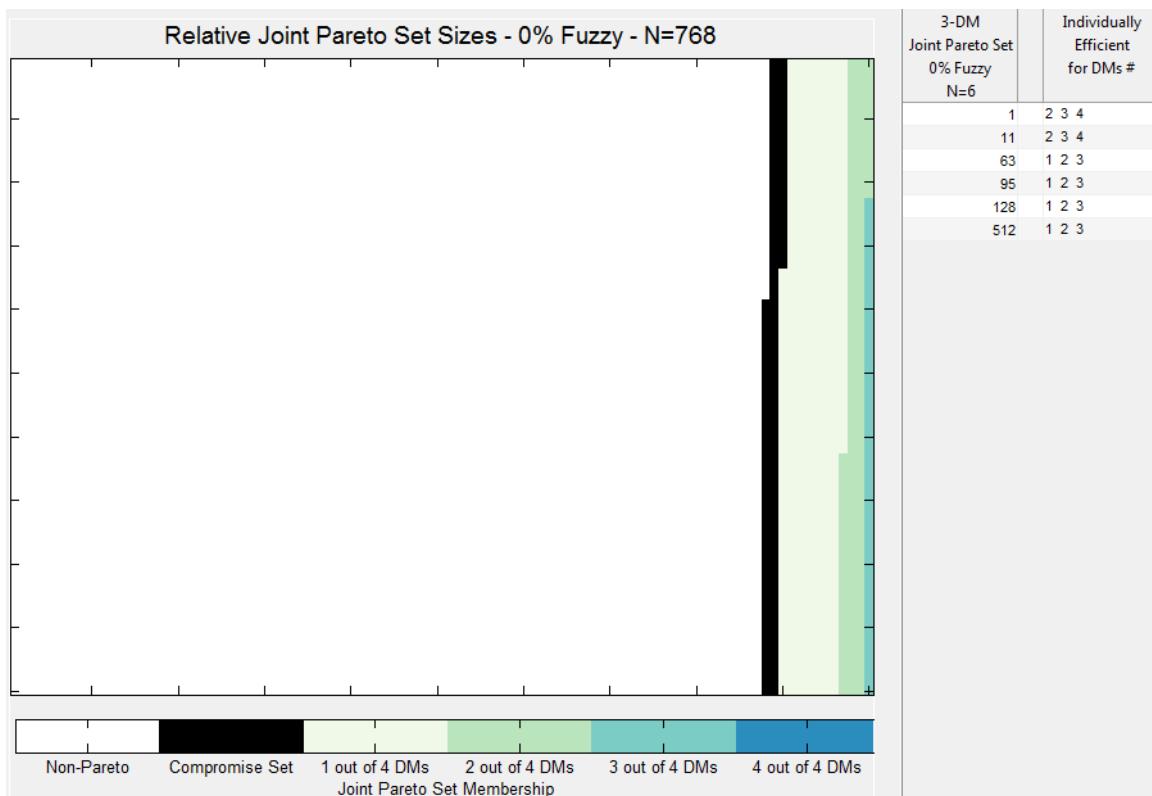
**Table 18.** Pareto sets sizes for value models using evaluative model #3

Objectives Set	# designs in 0% Pareto Set
1: MAU-COST	15
2: AHP-COST	54
3: CBA-COST	25
4: MOE-COST	20
JOINT	0
COMPROMISE	15

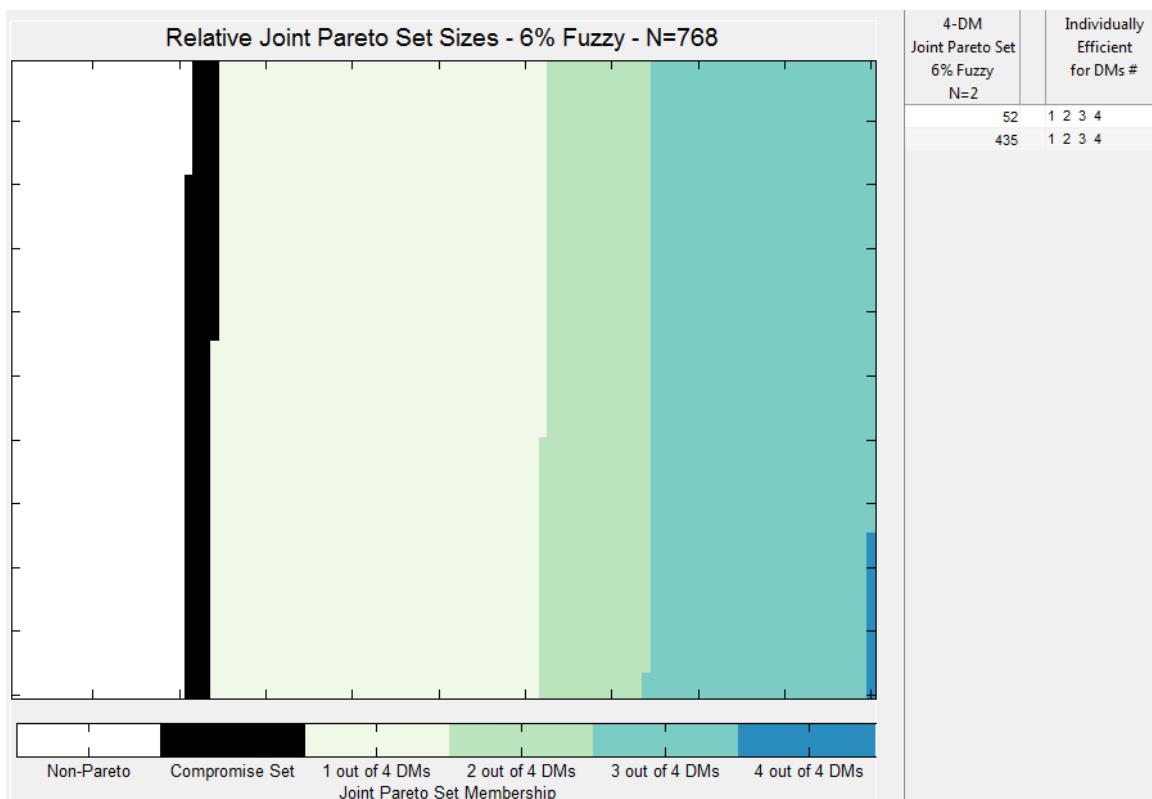
**Table 19.** Promising designs that are joint Pareto efficient across 3 out of 4 value models in evaluative model #3

ID Number	Pareto Efficient For	Invalid For	Details
1	2, 3, 4	1	small biprop min fuel
11	2, 3, 4	1	small cryo near min fuel
63	1, 2, 3		med nuke near max fuel
95	1, 2, 3		large nuke near max fuel
128	1, 2, 3		xl nuke max fuel
512	1, 2, 3		xl nuke max fuel CARBON

Two fully joint (across 4 out of 4 value models) Pareto efficient designs appears at a fuzzy level of 6%: design 52 (medium payload, electric propulsion, medium fuel tank, aluminum design), and design 435 (like 52, but with carbon and one size smaller fuel tank).



**Figure 33.** Gridmap showing relative sizes of 0% fuzzy joint Pareto sets in evaluative model #3



**Figure 34.** Gridmap showing relative sizes of 6% fuzzy joint Pareto sets in evaluative model #3

Figure 33 and Figure 34 illustrate the relative sizes of the fuzzy joint Pareto sets at 0% and 6% fuzzy levels respectively.

#### **Implementation #4 (new speed and material model)**

Recall evaluative model #4 has a potential tradespace size of 768 designs, enumerated across four design variables (payload size, propulsion type, fuel tank size, and material type). This model implementation incorporates both the (higher performance fidelity) new speed model as well as the (higher design fidelity) new material model. Table 20 describes the size of the 0% fuzzy Pareto sets for each of the value models when using evaluative model #4 (the new speed and material model). Table 21 lists the designs that are almost joint Pareto efficient, appearing in 3 out of 4 Pareto sets.

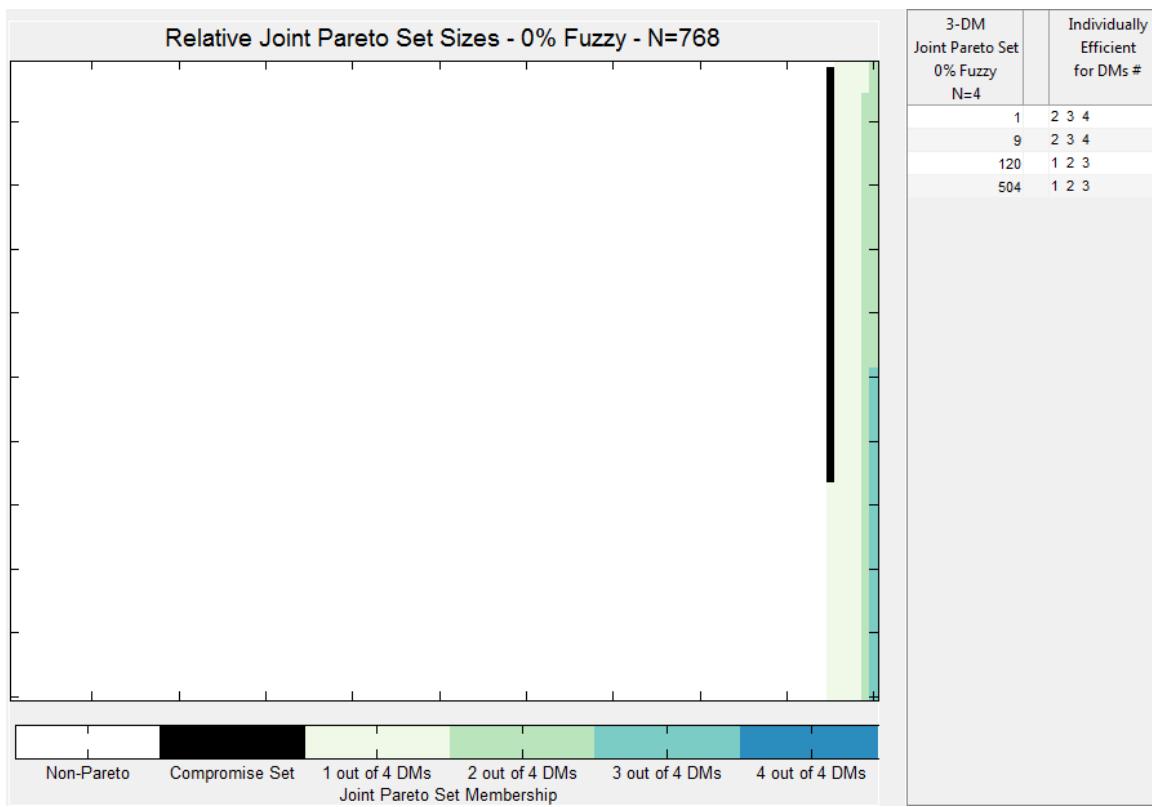
**Table 20. Pareto sets sizes for value models using evaluative model #4**

Objectives Set	# designs in 0% Pareto Set
1: MAU-COST	16
2: AHP-COST	19
3: CBA-COST	5
4: MOE-COST	20
JOINT	0
COMPROMISE	5

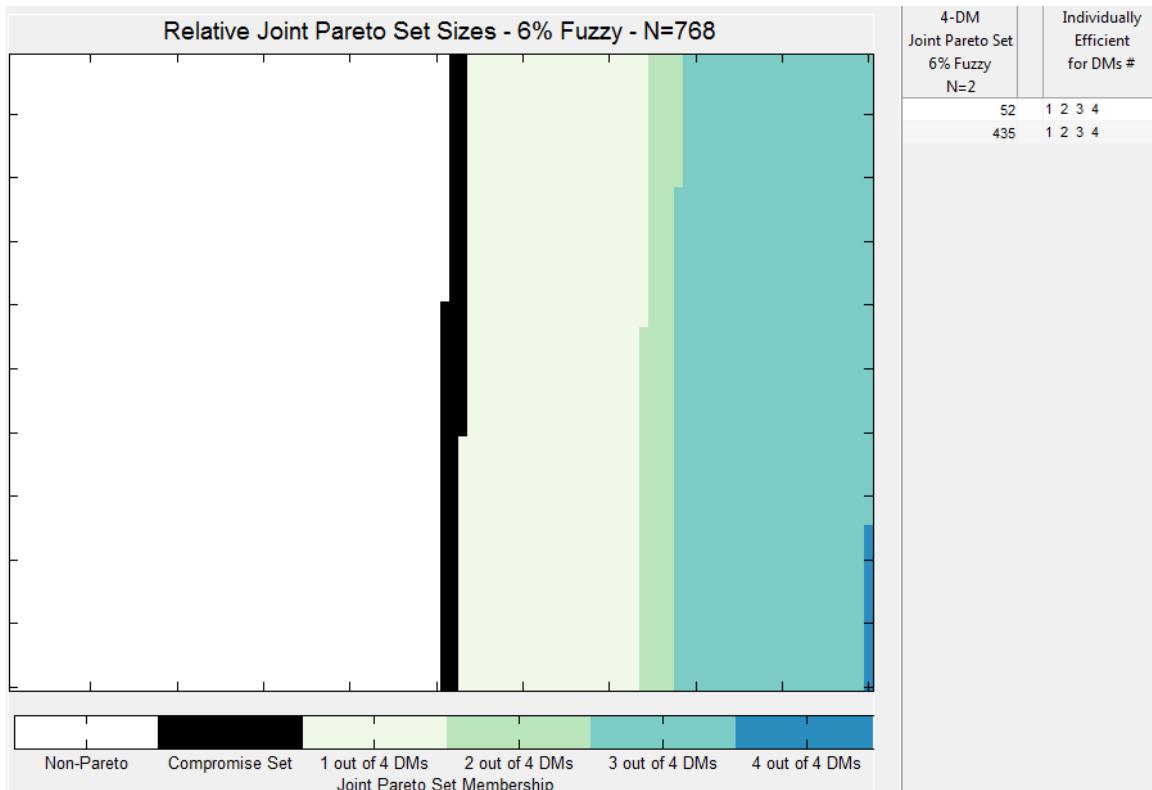
**Table 21. Promising designs that are joint Pareto efficient across 3 out of 4 value models in evaluative model #4**

ID Number	Pareto Efficient For	Invalid For	Details
1	2, 3, 4	1	small biprop min fuel
9	2, 3, 4	1	small cryo min fuel
120	1, 2, 3		xl elec max fuel
504	1, 2, 3		xl elec max fuel CARBON

The same two fully joint Pareto efficient designs appear at fuzzy level 6% as in evaluative model implementation #3: designs 52 and 435.



**Figure 35.** Gridmap showing relative sizes of 0% fuzzy joint Pareto sets in evaluative model #4



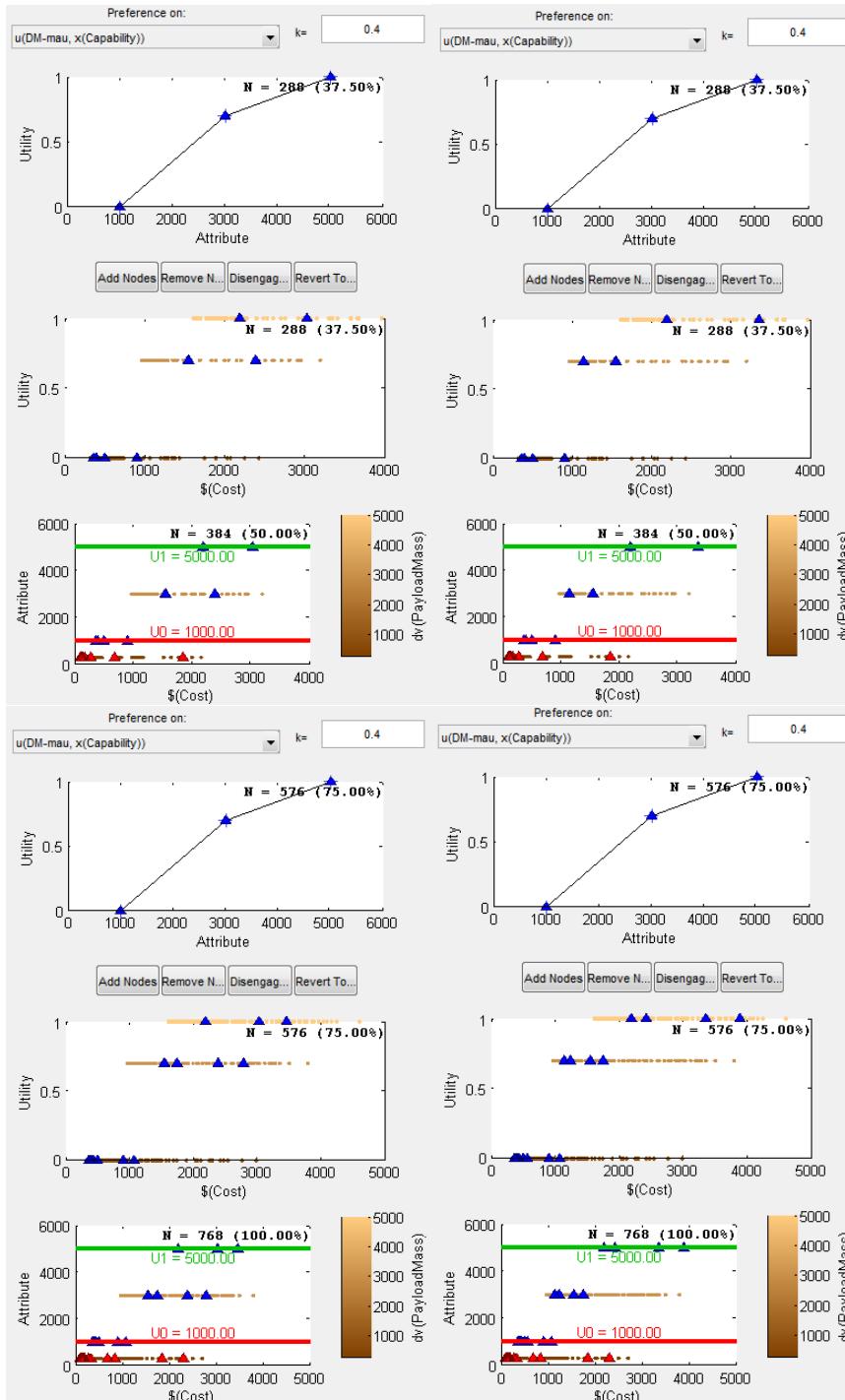
**Figure 36.** Gridmap showing relative sizes of 6% fuzzy joint Pareto sets in evaluative model #4

Figure 35 and Figure 36 illustrate the relative sizes of the fuzzy joint Pareto sets at 0% and 6% fuzzy levels respectively.

The addition of having carbon as a material type affects the first joint Pareto design, as this matches model #3. But in that case the change from 7% to 6% is likely an artifact of the increased cost range due to the addition of very expensive carbon designs to the tradespace (i.e. fuzzy level is calculated as a fraction of the tradespace cost range, so the addition of more expensive designs will make a given design appear relatively closer to the Pareto front). The new speed model makes electric more promising than nuclear type propulsion, as both model #2 and model #4 drop all promising nuclear designs and add promising electric designs.

### **Summary across the evaluative models**

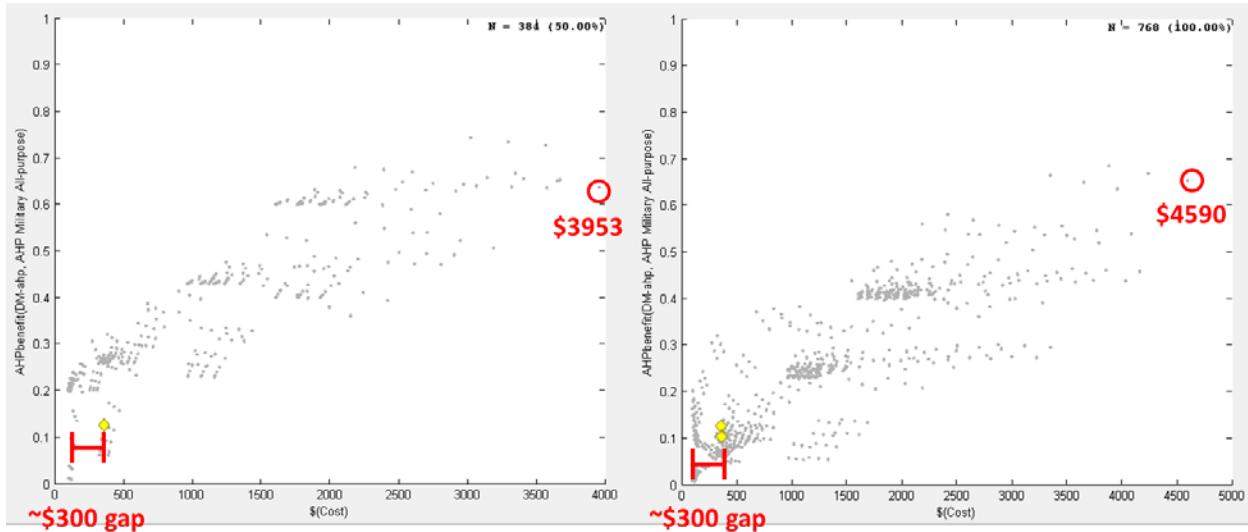
All four evaluative model implementations corroborate a tension between the MAU and MOE value models. That is, MOE value model prefers inexpensive, small designs in order to maximize delta-V, but none of those designs meet minimum acceptable performance levels for MAU, so they are considered invalid for the MAU value model.



**Figure 37. Comparison of the payload capability single attribute utility function impact across the four evaluative models with MAU (blue triangle) and MOE (red triangle) Pareto sets indicated**

As mentioned earlier, the MAU and MOE value models are in tension, as the MOE value model prefers low mass designs (strongly driven by small payloads). This can be seen in Figure 37, which illustrates the single attribute utility score as a function of payload capability across the four evaluative models. As payload capability goes up, so too does the mass of the alternatives. The red triangles indicate the designs that MOE views as Pareto efficient, while the blue triangles are

the designs in the MAU Pareto set. Many red triangles are below the U0 (red) line in the figure, which correspond to the minimum acceptable level of capability for the utility model. The MAU model (which includes the illustrated single attribute utility curves, as well as utility curves for other attributes), requires that payload capability be greater than the minimum possible size. This clearly shows that MOE and MAU value is in tension. This insight applies across all of the evaluative models and is an *insight about the nature of the value models*.



**Figure 38. Comparing evaluative model #1 first joint Pareto design at fuzzy level 7% (left) and evaluative model #4 first joint Pareto at fuzzy level 6% (right) illustrates how increase in maximum cost design impacts fuzzy metric**

Another insight described above is that the minimum fuzzy level for the appearance of joint Pareto designs decreased from 7% to 6% when adding the new material model (e.g. in implementations #3 and #4). However, design 52 is the design in both cases (with design 435 also in model #3 and #4), and according to the analysis, these designs do not move much across the evaluative models. Instead, the change in fuzzy level required appears to be an artifact of both how the fuzzy Pareto number is calculated, as well as the consequence of having more expensive designs appearing in the tradespace due to the new more expensive carbon version designs in models #3 and #4. This is an example of a *modeling artifact* and not an insight about the designs themselves.

### Evaluative Model Trading for each Value Model

For this part of the case, we look at the tradespaces for a given value model across the four different model implementations (essentially repeating the early evaluative model trade case for each of the four value models).

For the Pareto set category tables, the categories are different than in the corresponding Pareto sets described in the earlier evaluative trading case. The categories are consistent across the value model examples below, however. These categories now range from A to K, with each corresponding to different patterns in which models a design is Pareto efficient.

### Multi-attribute Utility (MAU)

Looking across the evaluative models within the MAU value model is what was done in the study above (evaluative model trading), so this section is abbreviated, with only Table 22 repeated below for easy reference. Six categories of designs are apparent in this table.

**Table 22. Designs, marked in gray with check, for model implementations in which they are efficient for the MAU value model**

Category	Design ID (Aluminum)	Model 1	Model 2	Model 3	Model 4	Design ID (Carbon)	Model 3	Model 4
A	52	✓	✓	✓	✓	436	✓	✓
	53	✓	✓	✓	✓	437	✓	✓
	63	✓	✓	✓	✓	447	✓	✓
B	54	✓	✓	✓	✓	438		✓
	87	✓	✓	✓	✓	471		✓
	119	✓	✓	✓	✓	503		✓
D	86		✓		✓	470		✓
	120		✓		✓	504		✓
E	96	✓		✓		480	✓	
	128	✓		✓		512	✓	
F	127	✓				511		
K	95	✓	✓	✓		479	✓	

- **A.** Designs 52, 53, and 63 (and their carbon counterparts) are always efficient across the model implementations.
- **B.** Designs 54, 87, and 119 are efficient across the model implementations, and their carbon counterparts are only efficient under the new speed model (model #4).
- **D.** Designs 86 and 120 are efficient only under the new speed model (models #2 and #4), as are their carbon counterparts (in model #4).
- **E.** Designs 96 and 128 are efficient only under the old speed model (models #1 and #3), as are their carbon counterparts (in model #3).
- **F.** Design 127 is efficient only under the original model (model #1).
- **K.** Design 95 (and its carbon counterpart) is efficient under each of the models except the combined model (model #4).

### Analytic Hierarchy Process (AHP)

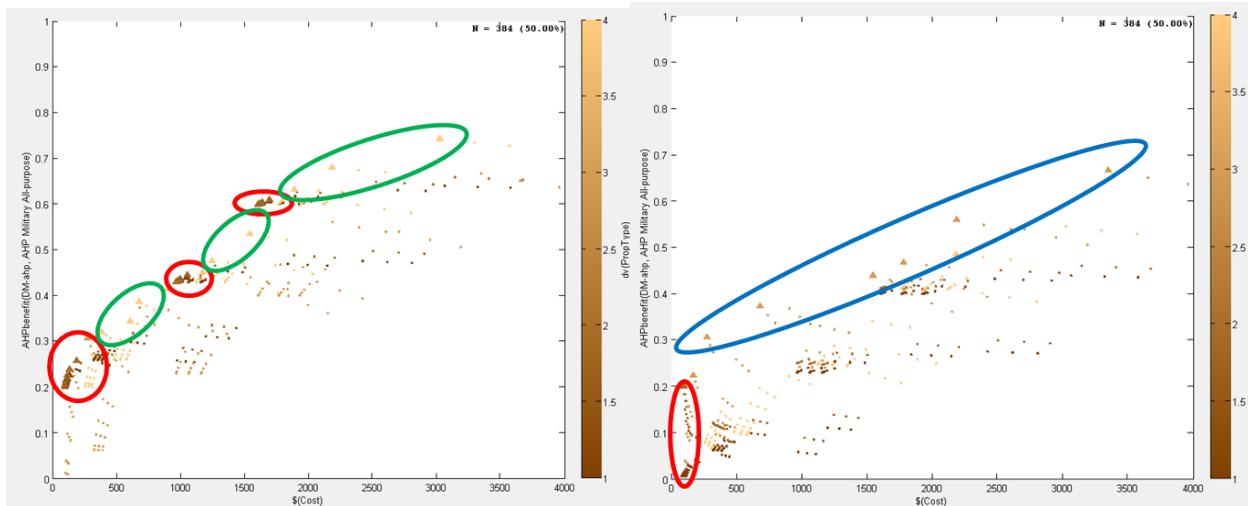
The next value model, AHP, was then compared across the evaluative models. First we identified designs that are in the Pareto set within each model. These designs are described in Table 23. Ten categories of designs are apparent in this table.

**Table 23. Designs, marked in gray with check, for model implementations in which they are efficient for the AHP value model**

Category	Design ID (Aluminum)	Model 1	Model 2	Model 3	Model 4	Design ID (Carbon)	Model 3	Model 4
A	22	✓	✓	✓	✓	406	✓	✓
B	9	✓	✓	✓	✓	393		✓
C	127	✓	✓	✓	✓	511	✓	
D	21		✓		✓	405		✓
	23		✓		✓	407		✓
	87		✓		✓	471		✓
	118		✓		✓	502		✓
	119		✓		✓	503		✓
	120		✓		✓	504		✓
E	10	✓		✓		394	✓	
	11	✓		✓		395	✓	
	12	✓		✓		396	✓	
	13	✓		✓		397	✓	
	14	✓		✓		398	✓	
	31	✓		✓		415	✓	
	77	✓		✓		461	✓	
	94	✓		✓		478	✓	
	109	✓		✓		493	✓	
	126	✓		✓		510	✓	
	128	✓		✓		512	✓	
F	4	✓				388		
	5	✓				389		
G	55					439		✓
H	1	✓	✓	✓	✓	385		
I	30	✓				414	✓	
J	2	✓		✓		386		
	3	✓		✓		387		
	62	✓		✓		446		
	63	✓		✓		447		
	65	✓		✓		449		
	66	✓		✓		450		
	67	✓		✓		451		
	68	✓		✓		452		
	69	✓		✓		453		
	73	✓		✓		457		
	74	✓		✓		458		
	75	✓		✓		459		
	76	✓		✓		460		

Category	Design ID (Aluminum)	Model 1	Model 2	Model 3	Model 4	Design ID (Carbon)	Model 3	Model 4
	93	✓		✓		477		
	95	✓		✓		479		
	97	✓		✓		481		
	98	✓		✓		482		
	99	✓		✓		483		
	100	✓		✓		484		
	101	✓		✓		485		
	105	✓		✓		489		
	106	✓		✓		490		
	107	✓		✓		491		
	108	✓		✓		492		
	125	✓		✓		509		

- **A.** Design 22 (and its carbon counterpart) is always efficient across the model implementations.
- **B.** Design 9 is efficient across the model implementations, and its carbon counterpart (design 393) is only efficient under the new speed model (model #4).
- **C.** Design 127 is efficient across the model implementations, and its carbon counterpart (design 511) is only efficient under the old speed model (model #3).
- **D.** Designs 21, 23, 87, 118, 119, and 120 are efficient only under the new speed model (models #2 and #4), as are their carbon counterparts (in model #4).
- **E.** Designs 10, 11, 12, 13, 14, 31, 77, 94, 109, 126, and 128 are efficient only under the old speed model (models #1 and #3), as are their carbon counterparts (in model #3).
- **F.** Designs 4 and 5 are efficient only under the original model (model #1).
- **G.** Design 439 is the carbon version of design 55, but is only efficient in the carbon variant under the new speed model (model #4).
- **H.** Design 1 is efficient across all of the model implementations, but its carbon variant is never efficient.
- **I.** Design 30 is only efficient in the original model (model #1) and its carbon counterpart replaces it as efficient under the old speed model (model #3).
- **J.** These designs are only efficient in the old speed model (model #1 and model #3), but their carbon variants are never efficient.



**Figure 39.** AHP tradespace scatterplot under the old speed models (#1 and #3, left) and new speed models (#2 and #4, right), with families of designs indicated by propulsion type (red=biprop/cryo, green=nuclear, blue=electric)

As can be seen in Figure 39, electric vehicles replace nuclear vehicles on the Pareto front under the new speed model (old speed models #1 and #3 on left, and new speed models #2 and #4 on right).

#### Cost-Benefit Analysis

The next value model, CBA, was then compared across the evaluative models. First we identified designs that are in the Pareto set within each model. These designs are described in Table 24. Six categories of designs are apparent in this table.

**Table 24.** Designs, marked in gray with check, for model implementations in which they are efficient for the CBA value model

Category	Design ID (Aluminum)	Model 1	Model 2	Model 3	Model 4	Design ID (Carbon)	Model 3	Model 4
D	9		✓		✓	393		✓
	120		✓		✓	504		✓
E	11	✓		✓		395	✓	
	12	✓		✓		396	✓	
	13	✓		✓		397	✓	
	14	✓		✓		398	✓	
	29	✓		✓		413	✓	
	30	✓		✓		414	✓	
	31	✓		✓		415	✓	
	63	✓		✓		447	✓	
	95	✓		✓		479	✓	
	96	✓		✓		480	✓	
F	128	✓		✓		512	✓	
	4	✓				388		
	5	✓				389		

Category	Design ID (Aluminum)	Model 1	Model 2	Model 3	Model 4	Design ID (Carbon)	Model 3	Model 4
	127	✓				511		
H	1	✓	✓	✓	✓	385		
J	3	✓		✓		387		
	32	✓		✓		416		
L	87		✓			471		
	88		✓			472		
	119		✓			503		

- D. Designs 9 and 120 are efficient only under the new speed model (models #2 and #4), as are their carbon counterparts (in model #4).
- E. Designs 11, 12, 13, 14, 29, 30, 31, 63, 95, 96, and 128 are efficient only under the old speed model (models #1 and #3), as are their carbon counterparts (in model #3).
- F. Designs 4, 5, and 127 are efficient only under the original model (model #1).
- H. Design 1 is efficient across all of the model implementations, but its carbon variant is never efficient.
- J. These designs are only efficient in the old speed model (model #1 and model #3), but their carbon variants are never efficient.
- L. Designs 87, 88, and 119 are only efficient in the new speed model (model #2) without carbon as an option (model #3)

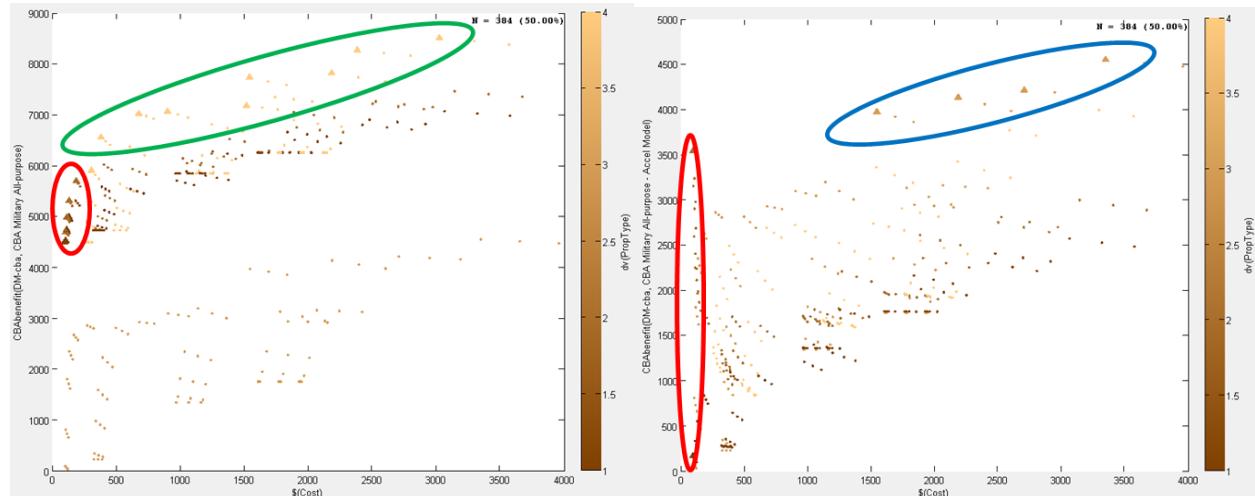


Figure 40. CBA tradespace scatterplot under the old speed models (#1 and #3, left) and new speed models (#2 and #4, right), with families of designs indicated by propulsion type (red=biprop/cryo, green=nuclear, blue=electric)

Similar to the pattern for the AHP value model, as can be seen in Figure 40, electric vehicles replace nuclear on the Pareto front under the new speed model.

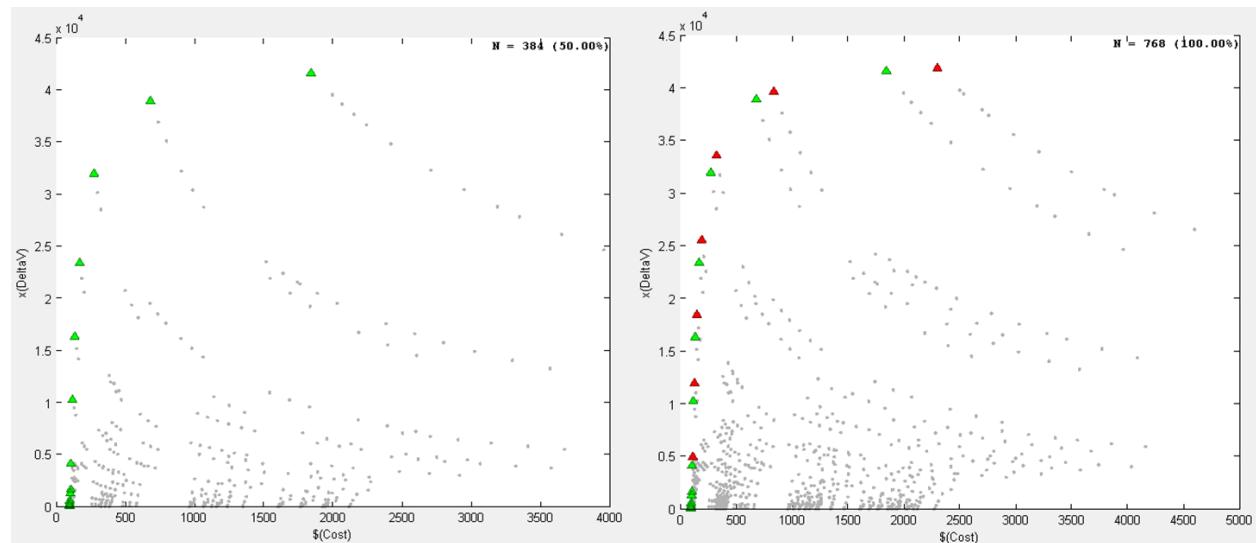
### Measure of Effectiveness

The next value model, MOE, was then compared across the evaluative models. First we identified designs that are in the Pareto set within each model. These designs are described in Table 25. Two categories of designs are apparent in this table.

**Table 25. Designs, marked in gray with check, for model implementations in which they are efficient for the MOE value model**

Category	Design ID (Aluminum)	Model 1	Model 2	Model 3	Model 4	Design ID (Carbon)	Model 3	Model 4
A	18	✓	✓	✓	✓	402	✓	✓
	19	✓	✓	✓	✓	403	✓	✓
	20	✓	✓	✓	✓	404	✓	✓
	21	✓	✓	✓	✓	405	✓	✓
	22	✓	✓	✓	✓	406	✓	✓
	23	✓	✓	✓	✓	407	✓	✓
	24	✓	✓	✓	✓	408	✓	✓
H	1	✓	✓	✓	✓	385		
	2	✓	✓	✓	✓	386		
	9	✓	✓	✓	✓	393		
	10	✓	✓	✓	✓	394		
	11	✓	✓	✓	✓	395		
	17	✓	✓	✓	✓	401		

- **A.** Designs 18, 19, 20, 21, 22, 23, and 24 (and their carbon counterparts) are always efficient across the model implementations.
- **H.** Designs 1, 2, 9, 10, 11, and 17 are efficient across all of the model implementations, but their carbon variants are never efficient.



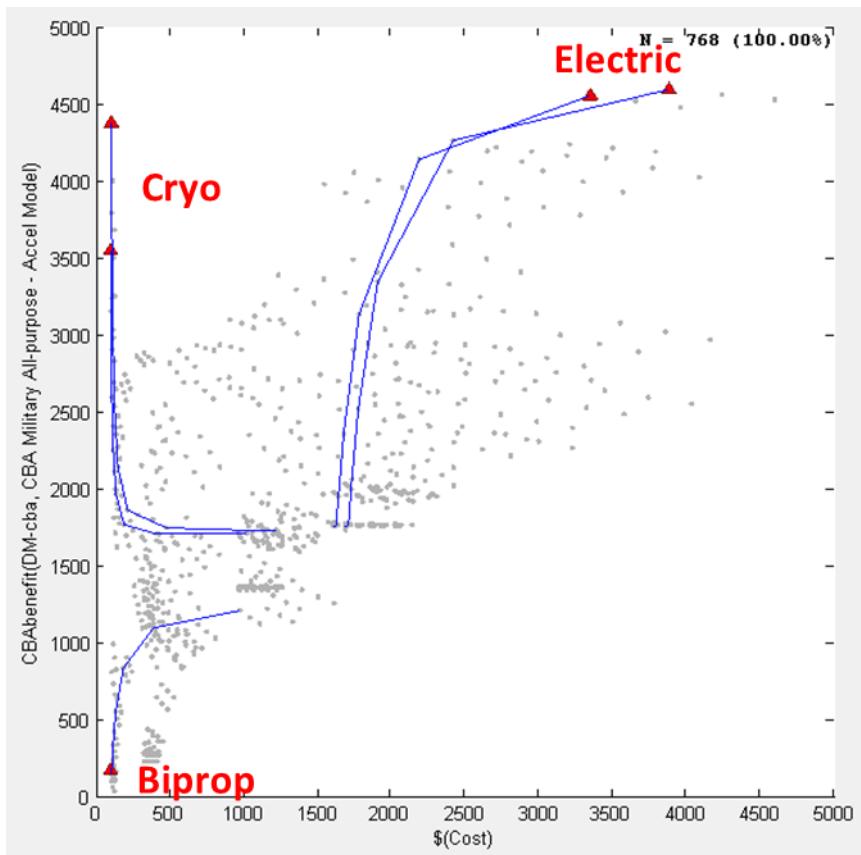
**Figure 41. Comparison of MOE value model across evaluative models #1 and #2 (left) and models #3 and #4 (right)**

Figure 41 shows the Pareto set in evaluative models #1 and #2 (left) in green triangles, and in models #3 and #4 (right) with the red triangles corresponding to the new carbon-version designs added in models #3 and #4. Notice how the sets do not change much, as the red points are just minor cost-increased version of their aluminum counterparts in green. This is due to the fact that this MOE ( $\Delta V$ ) is unaffected by the change in the model implementations relative to the original model. Carbon becomes Pareto efficient when cost gets high enough (i.e. not for the first couple of small designs). Small payloads dominate  $\Delta V$ .

### Discussion of Combined Value and Evaluative Model Case

The key take away from the combined value and evaluative model case is that doing this kind of activity can reveal two classes of insights:

- 1) insights into fundamental relationship between perspectives of value and what is possible (*structural patterns for the decision problem*), and
- 2) insights into modeling artifacts, both in how value is capture and how evaluation is performed (*modeling artifacts*).



**Figure 42. CBA value model scatterplot for evaluative model #4 (combined new material and new speed)**

The first class of insights sometimes appear to emerge through the course of analysis. This may be due to the fact that the relationships are buried in the interactions between factors of the problem and are not readily apparent in our mental models. For example, Figure 42 below shows

the CBA value model scatterplot in the evaluative model #4 (combined new material and new speed model). Lines show Pareto efficient points connected to similar designs with different levels of fuel. For both the electric and biprop propulsion types there is a positive tradeoff of more fuel (= more cost) for more benefit. But counterintuitively the small cryo propulsion designs actually get less benefit for more fuel. This is because the added fuel actually decreases the speed of those small designs in spite of increasing the on-board delta-V. This is a consequence of the confluence of physics (i.e. the rocket equation and inertia) and expectations on what is considered beneficial. The very fact that the relationship between fuel mass and benefit plays out differently in different regions of the tradespace means that the complexity (in terms of number of factors to consider) likely would overwhelm an unaided human mind due to bounded rationality.

Each evaluative model is one representation of how a system might “perform”, while each value model is one representation of how a system might “be valued.” The emergence described above would occur at the intersection of each possible evaluative and value model, as well as across them, as shown in this simple Space Tug demonstration. Systems engineers and analysts may benefit strongly by considering not only their choice of evaluative and value model, but also how their insights might vary if they were to include more than one of each type of model.

---

#### NEXT STEPS

Future work will synthesize the phase 3 results into a more prescriptive framework. The combined value and evaluative model results will be published. The efforts of this work have been used to identify additional needs for research in the IMCSE research roadmap. Related areas of investigation will be pursued under the next phase activity on curation of model-centric environments.

## SYNTHESIS OF IMCSE RESEARCH PROGRESS

IMCSE research progress has been made in a number of key areas, performing consequential research in support of DoD imperatives.

The IMCSE Pathfinder project has seeded a shared research vision and an initial research roadmap. This is aligned with the desire to investigate core principles of engineering and science through collaborative efforts.

DoD's need to create and respond to the dynamic digital model-centric ecosystem is supported by IMCSE investigations on cognitive and perceptual aspects of human-model interaction, and by investigation of multi-stakeholder negotiation.

Baldwin, K. Defense Engineering Excellence, NDIA, Oct 26-29, 2015

IMCSE RESEARCH

**IMCSE Pathfinder**  
engage community in  
creating a shared  
research vision and  
collaborative roadmap

"We need to investigate the core principles of engineering & science that lay the foundation for significant, next generation advances in cross-discipline engineering practice and education in multi-scale environments." IAWG Joint Statement

Cognitive and Perceptual Aspects of Human-Model Interaction  
Empirical case studies of transitions in interaction

Multi-Stakeholder Negotiation  
Human to human interaction via models

Future: Dynamic Digital Model-Centric Ecosystem  
Zimmerman, P., A Framework for Developing a Digital System Model Taxonomy, NDIA, Oct 27, 2015

In support of the vision for digital system models and modular/open systems, IMCSE investigation has explored several key topics. A study on the non-technical challenges for digital system models explored issues and possible structures through the technology-policy lense.

The ability to transition proprietary model implementations to open source was demonstrated by the Interactive Schedule Reduction model prototype. The many facets of strategic management of model-centric environments was explored in a preliminary study on the considerations for model curation.

- Digital System Model/Digital Thread
- Education
- Policy & Guidance
- Data Rights

**IMCSE RESEARCH**

**Digital System Models: Investigating the Non-Technical Challenges**  
*exploration of issues and possible structures through the technology-policy lens*

**Interactive Schedule Reduction Model Prototype**  
*demonstration prototype, method/proprietary model transition to open source*

**Consideration for Model Curation**  
*understanding curator role, curation function, and constructs and standards*

Zimmerman, P., Modularity and Open Systems: Meaningful Distinctions, NDIA, Oct 27, 2015

**IEEA Framework**

- The process of IEEA aim to address the need for early-stage conceptual modeling of systems with sustainability in mind.
- Modifying RCM/IMACS to explicitly include iterative and adaptive modeling considerations enables the user to make higher-level decisions impact “value” management.

IMCSE has made progress on research supporting Engineered Resilient Systems (ERS). This includes a framework for interactive Epoch-Era Analysis, transforming an innovated method with interactive model-centric capability. Interactive visualization tools and analytics were investigated and demonstration prototypes and cases provide evidence of new constructs and tools that are grounded in the science of visual analytics. The importance of model choice and tradeoff was investigated and demonstrated in application cases.

**Engineered Resilient System**

Developing integrated suite of modern engineering tools, models and related capabilities, tradespace assessment and visualization tools; all within an architecture aligned with acquisition and operational business processes.

Zimmerman, P., Modularity and Open Systems: Meaningful Distinctions, NDIA, Oct 27, 2015

**IMCSE RESEARCH**

**Interactive Epoch-Era Analysis**  
*transforming an innovative method with interactive model-centric capability*

**Interactive Visualization Tools and Analytics**  
*demonstration prototypes and case application of new constructs and tools grounded in science of visual analytics*

**Model Choice and Tradeoff**  
*Enrich model use and decision making through model tradeoff*

Results of the IMCSE research progress are discussed in a series of publications.

## **2014/2015**

- Fitzgerald, M.E., and Ross, A.M., "Controlling for Framing Effects in Multi-Stakeholder Tradespace Exploration," 12th Conference on Systems Engineering Research, Redondo Beach, CA, March 2014 (SERC STUDENT PAPER AWARD)
- Curry, M, and Ross, A.M., "Considerations for an Extended Framework for Interactive Epoch-Era Analysis," 13th Conference on Systems Engineering Research, Hoboken, NJ, March 2015.
- Fitzgerald, M.E., and Ross, A.M., "Effects of Enhanced Multi-party Tradespace Visualization on a Two-person Negotiation," 13th Conference on Systems Engineering Research, Hoboken, NJ, March 2015
- Grogan, P.T., de Weck, O.L., Ross, A.M., and Rhodes, D.H., "Interactive Models as a System Design Tool: Applications to System Project Management," 13th Conference on Systems Engineering Research, Hoboken, NJ, March 2015
- Ross, A.M., Rhodes, D.H., and Fitzgerald, M.E., "Interactive Value Model Trading for Resilient Systems Decisions," 13th Conference on Systems Engineering Research, Hoboken, NJ, March 2015 (CSER BEST ACADEMIC PAPER AWARD)

## **2016**

- Curry, M, and Ross, A.M., "Designing System Value Sustainment using Interactive Epoch Era Analysis: A Case Study for On-orbit Servicing Vehicles," 14th Conference on Systems Engineering Research, Huntsville, AL, March 2016.
- German, E.S., and Rhodes, D.H., "Human-Model Interactivity: What Can Be Learned from the Experience of Pilots Transitioning to Glass Cockpit?," 14th Conference on Systems Engineering Research, Huntsville, AL, March 2016
- Reid, J.B., and Rhodes, D.H., "Digital System Models: An Investigation of the Non-technical Challenges and Research Needs," 14th Conference on Systems Engineering Research, Huntsville, AL, March 2016
- Ross, A.M., Fitzgerald, M.E., and Rhodes, D.H., "Interactive Evaluative Model Trading for Resilient Systems Decisions," 14th Conference on Systems Engineering Research, Huntsville, AL, March 2016
- Reymondet, L., Rhodes, D.H., and Ross, A.M., "Considerations for Model Curation in Model-Centric Systems Engineering," 10th Annual IEEE Systems Conference, Orlando, FL, April 2016
- Fitzgerald, M.E., and Ross, A.M., "Recommendations for Framing Multi-Stakeholder Tradespace Exploration," INCOSE International Symposium 2016, Edinburgh, Scotland, July 2016.
- Rhodes, D.H., and Ross, A.M., "A Vision for Human-Model Interaction in Interactive Model-Centric Systems Engineering," INCOSE International Symposium 2016, Edinburgh, Scotland, July 2016

## MOVING FORWARD TO PHASE FOUR

---

IMCSE advances the current state of SE knowledge in “non-technical” aspects of model-based engineering. While MBSE and MBE activities are advancing technical aspects of models in the engineering of systems, this topic advances knowledge relevant to human interaction with models and model-generated information. The next phase of the research is expected to further contribute to addressing the SERC’s expressed need for research in support of the ability to successfully build adaptive, resilient and secure complex defense systems that are increasingly constrained by how well those responsible for conceiving, developing, manufacturing, and operating those systems can work together effectively. This ability depends upon the effective interaction of humans in model-centric environments, which remains a multi-faceted challenge.

Phase 4 addresses a subset of the research needs in the emerging research agenda for interactive model-centric systems engineering, focusing on three activities:

1. **Interactive Epoch-Era Analysis (IEEA).** The researchers will continue to mature the approach for evaluating systems under dynamic uncertainty, with further development of the extended framework for interactive capability and scaling to big data. New research includes enhancing interactive capability, investigating additional visual analytic techniques and tools, and applying the maturing prototype framework with supporting tools in additional case analysis and application. An impact assessment for use of IEEA methods will be developed.
2. **Human-Model Interaction.** The prior phase confirmed the need for empirical investigation of the needs and desires for the interactive experience from the perspective of users of model-centric environments. The researchers will investigate interactions of individuals with models, and multi-stakeholder teams using model-generated information. Additionally, the researchers will continue investigating the analogy case to characterize the underlying cognitive considerations in human-model interaction, with the intent of developing heuristics/design principles for trial use by model developers, model users, and model-based software designers, leading to a validated set of guiding principles for effective human-model interaction.
3. **Curation of Interactive Model-Centric Environments.** Interaction with government and industry leadership and practitioners in the prior phase revealed a number of “non-technical” research needs in support of sustainable model-centric environments designed for highly effective human engagement with formal practices for managing models as trusted assets. The researchers will investigate the needs and benefits of formal curation of model-centric environments. Preliminary investigation of the topic revealed the need for enhancing human interactivity with models and managing models as trusted assets, as well as foundational taxonomies and capabilities. Curation roles and responsibilities, and potential impacts on current processes will be elicited and further developed. The expected outcome is a generic roles & responsibilities document for a curation function in interactive model-centric environments, grounded in empirical findings.

## REFERENCES

---

- Aisenberg, M.A., Davis, D., Kehoe, T.D., "Intellectual Property Rights for Digital Design and Manufacturing: Issues and Recommendations", The MITRE Corporation, 2014.
- Alexander,A. Brunyé, T., Sidman, J., Weil S.A. "From gaming to training: A review of studies on fidelity, immersion, presence, and buy-in and their effects on transfer in pc-based simulations and games". DARWARS Training Impact Group 2005, vol. 5, 1-14.
- Bankes, S., "Exploratory Modeling for Policy Analysis", in Operations Research, Vol. 41, No. 3, 1993.
- Baxter, G., Besnard, D. and Riley, D. "Cognitive mismatches in the cockpit: Will they ever be a thing of the past?" Applied Ergonomics, 2010: 417-423.
- Besnard, D. and Baxter, G. "Cognitive Conflicts in Dynamic Systems." In Structure for Dependability: Computer-Based Systems from an Interdisciplinary Perspective, 107-124. London: Springer London, 2006.
- Blackburn, M., R. Cloutier, E. Hole, M. Bone, Transforming System Engineering through Model-Centric Engineering, Technical Report SERC-2015-TR-044-3, 2015.
- Chappell, Alan R, Edward G Crowther, Christine M Mitchell, and T Govindaraj. "The VNAV Tutor: Addressing a Mode Awareness Difficulty for Pilots of Glass Cockpit Aircraft." IEEE Transactions on Systems, Man and Cybernetics - Part A: Systems and Humans , 1997: 372-385.
- Cornell University, Trade Secret, WEX. Legal Information Institute, Cornell Law School, Ithaca, NY.
- Curry, M, and Ross, A.M., "Considerations for an Extended Framework for Interactive Epoch-Era Analysis," 13th Conference on Systems Engineering Research, Hoboken, NJ, March 2015.
- Digital Tapestry, Lockheed Martin. Available: <http://www.lockheedmartin.com/us/what-we-do/emerging/advanced-manufacturing/digital-tapestry.html> [Accessed: 10/15/2015]
- DoD, Systems Engineering Guide for Systems of Systems. Systems and Software Engineering, Department of Defense, 2008.
- DoD, Glossary: Defense Acquisition Acronyms and Terms. Department of Defense, US Department of Defense, Defense Acquisition University, Learning Capabilities Integration Center, Center for Acquisition and Program Management, 2015.
- Endsley, Mica R. "Automation and Situation Awareness." In Automation and Human Performance: Theory and Applications, 163-181. Mahwah: Lawrence Erlbaum, 1996.
- FDA, The Sentinel Initiative. Office of Surveillane and Epidemiology, Federal Drug Administration, 2010
- Feeney, A. Digital Thread for Smart Manufacturing. Engineering Laboratory-Systems Integration Division-Systems Engineering Group, US National Institute of Standards and Technology, 2014.
- Fitzgerald, M., and Ross, A.M., "Effects of Enhanced Multi-party Tradespace Visualization on a Two-person Negotiation", Procedia Computer Science 2015; vol. 44, 466-475.
- Fitzgerald, M.E. and Ross, A.M., "Mitigating Contextual Uncertainties with Valuable Changeability Analysis in the Multi-Epoch Domain," 6th Annual IEEE Systems Conference, Vancouver, Canada, March 2012.
- Fitzgerald, M.E., Managing Uncertainty in Systems with a Valuation Approach for Strategic Changeability, Master of Science Thesis, Aeronautics and Astronautics, MIT, June 2012.
- Fulcoly, D.O., Ross, A.M., and Rhodes, D.H., "Evaluating System Change Options and Timing using the Epoch Syncopation Framework," 10th Conference on Systems Engineering Research, St. Louis, MO, March 2012.
- G. E. P. Box, Science and Statistics, Journal of the American Statistical Association, 1976.

- Gaspar, H., Rhodes, D.H., Ross, A.M., and Erikstad, E.O., "Addressing Complexity Aspects in Conceptual Ship Design: A Systems Engineering Approach" *Journal of Ship Production and Design*, Vol. 28, No. 4, Nov 2012, pp. 145-159.
- German, E.S. and Rhodes, D.H., "Human-Model Interactivity: What can be learned from the experience of pilots transitioning to glass cockpit?", Conference on Systems Engineering Research, March 2016.
- Glaessgen,E. and Stargel. D., The Digital Twin paradigm for future NASA and US Air Force vehicles. 53rd Struct. Dyn. Mater. Conf.; 2012 1-14.
- Goh, G. Building Singapore's 'digital twin'. Digital News Asia; 2015
- Heer, J. and Agrawala, M., "Design considerations for collaborative visual analytics," *Information Visualization*, 7(1):49–62, 2007.
- Heer, J., and Shneiderman, B., "Interactive Dynamics for Visual Analysis," *ACM Queue*, 2012.
- Hiremath, V. Proctor, R., Fanjoy, R., Feyen, R., and Young, J. "Comparison of Pilot Recovery and Response Times in Two Types of Cockpits." In *Human Interface and the Management of Information*, 766-775. San Diego: Springer Berlin Heidelberg, 2009.
- Icke, I. and Sklar, E., "Visual Analytics: A Multi-faceted Overview," NSF Technical Report, April 16, 2009.
- INCOSE, INCOSE SE Vision 2020 (INCOSE-TP-2004-004-02, Sep 2007).
- INCOSE, Systems Engineering Handbook: A Guide for system life cycle processes and activities. 4th ed. Chapter 9.1 Modeling and Simulation. INCOSE-TP-2003-002-04. 2015.
- Keim, D., Andrienko, G., Fekete, J., Gorg, C., Kohlhammer, J., and Melançon, G., "Visual analytics: Definition, process, and challenges," *Information Visualization*, pages 154–175. Springer, 2008.
- Keim. D., "Information visualization and visual data mining," *IEEE Trans. on Visualization and Computer Graphics*, 8(1):1–8, 2002.
- Kellner, T., Wind in the Cloud? How the Digital Wind Farm Will Make Wind Power 20 Percent More Efficient. GE Reports; 2015.
- Kolmogorov, A. N., "Combinatorial foundations of information theory and the calculus of probabilities, *Russian Mathematical Surveys*, 38, 4, 27–36, 1983.
- Kraft, E. M., HPCMP CREATE TM-AV and the Air Force Digital Thread. SciTech; 2015.
- Krolak, P., Felts, W., and Marble, G., "A Man-Machine Approach Toward Solving the Traveling Salesman Problem," *Communications of the ACM*, Volume 14, Number 5, May 1971.
- Lee BD, Binder WR, Paredis CJ., "A Systematic Method For Specifying Effective Value Models," CSER14. March 2014.
- Lockheed Martin, The Digital Thread Key to F35 Joint Strike Fighter Affordability, Lockheed Martin. Available: <https://www.onlineamd.com/article/amd-080910-f-35-joint-strike-fighter-digital-thread> [Accessed: 9/22/16].
- Mathieu, J. ,Melhuish, J. James, J., Mahoney, P. Boiney, L., White, B. "Multi-Scale Modeling of the Air and Space Operations Center", The MITRE Corporation, 2007.
- Mathieu, J. James, J. Mahoney, P. Boiney, L., Hubbard, R., White, B. "Hybrid Systems Dynamic, Petri Net, and Agent-Based Modeling of the Air and Space Operations Center", The MITRE Corporation, 2007
- McBurney, P., "What are Models for?", in: M. Cossentino, K. Tuyls and G. Weiss (Editors): Post-Proceedings of the Ninth European Workshop on Multi-Agent Systems (EUMAS 2011). Lecture Notes in Computer Science, volume 7541. Berlin, Germany: Springer, pages 175- 188.
- McGrath, M., "Protecting the Digital Thread". Global Supply Chain Summit; 2014.
- McManus H. Schuman T. "Understanding the orbital transfer vehicle trade space", AIAA Space 2003. Long Beach, CA, September 2003.

- Morgan, W. G., Fischhoff, B., Bostrom, A., Lave, L., and Atman, C. ES&T Features: Communicating Risk to the Public - First, Learn what people know and believe. *Environmental Science & Technology* 1992; vol. 26, 2048-2056.
- Mosier, Kathleen L, and Linda J Skitka. "Automation Use and Automation Bias." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Santa Monica: Human Factors and Ergonomics Society, 1999. 344-348
- Mosier, Kathleen L, Linda J Skitka, Melisa Dunbar, and Lori McDonnell. "Aircrues and Automation Bias: The Advantages of Teamwork?" *The International Journal of Aviation Psychology* 11, no. 1 (2001): 1-14
- Mosier, Kathleen L, Linda J Skitka, Susan Heers, and Mark D Burdick. "Automation Bias: Decision Making and Performance in High-Tech Cockpits." *The International Journal of Aviation Psychology*, 2009: 47-63
- Mosier, Kathleen L, Nikita Sethi, Shane McCauley, Len Khoo, and Judith M Orasanu. "What You Don't Know Can Hurt You: Factors Impacting Diagnosis in the Automated Cockpit." *Human Factors*, 2007: 300-317.
- NASA. Standard for Models and Simulations. NASA-STD-7009; July 2008.
- National Research Council (2007), Human-Systems Integration in the System Development Process: A New Look, Committee on Human-System Design Support for Changing Technology, R.W. Pew and A.S. Mavor, Eds., Committee on Human Factors, Division of Behavioral and Social Sciences and Education, Washington, DC: The National Academies Press.
- Nau, D.S., "Pathology on Game Trees Revisited, and an Alternative to Minimaxing," *Artificial Intelligence*, 221-244, 1983.
- NDIA, Final Report of Model Based Engineering (MBE) Subcommittee, NDIA Systems Engineering Division M&S Committee, National Defense Industrial Association, 2011.
- Nickerson, Raymond S. "Confirmation Bias: A Ubiquitous Phenomenon in Many Guises." *Review of General Psychology*, 1998: 175-220.
- NTSB, Introduction of Glass Cockpit Avionics into Light Aircraft. National Transportation Safety Board. March 9, 2010.
- Oliver, D., Kelliher, T., Keegan, J., *Engineering Complex Systems with Models and Objects*, NY: McGraw-Hill, 1997.
- Parasuraman, R. and Manzey, D., "Complacency and Bias in Human Use of Automation: An Attentional Integration." *Human Factors*, 2010: 381-410.
- Park, S. and Rothrock, L., "Systematic analysis of framing bias in missile defense: Implications toward visualization design". *European Journal of Operational Research* 2007; vol. 182, 1383-1398.
- Pronin, E., Lin, D. and Ross, L. "The Bias Blind Spot: Perceptions of Bias in Self Versus Others." *Per Soc Psychol Bull*, 2002: 369-381.
- Reid, J.B. and Rhodes, D.H., "Digital System Models: An investigation of the non-technical challenges and research needs", Conference on Systems Engineering Research, March 2016.
- Reymondet L, Rhodes DH, and Ross AM, "Considerations for Model Curation in Model-Centric Systems Engineering," IEEE SysCon 2016, Orlando, FL, April 2016.
- Rhodes DH, and Ross AM, "A Vision for Human-Model Interaction in Model-Centric Systems Engineering," INCOSE Int'l Symp 2016, Edinburgh, UK, July 2016.
- Rhodes, D.H. and Ross,A.M., Interactive Model-Centric Systems Engineering (IMCSE) Pathfinder Workshop Report, Systems Engineering Research Center, Systems Engineering Advancement Research Initiative; January 2015.
- Rhodes, D.H. and Ross,A.M., Interactive Model-Centric Systems Engineering Technical Report SERC-2014-TR 048-2, February 28, 2015.
- Rhodes DH, Ross AM. Interactive Model-Centric Systems Engineering (IMCSE) Phase 1 Technical Report. SERC-2014-TR-048-1; September 2014.

- Rhodes, D.H. and Ross, A.M., "Five aspects of engineering complex systems: emerging constructs and methods", Proceedings, 4th Annual IEEE Systems Conference, April, San Diego, CA, 2010.
- Rhodes, D.H. and Ross, A.M., "Anticipatory Capacity: Leveraging Model-Based Approaches to Design Systems for Dynamic Futures," 2nd Annual Conference on Model-based Systems, Haifa, Israel, March 2009.
- Ricci, N., Schaffner, M.A., Ross, A.M., Rhodes, D.H., and Fitzgerald, M.E., "Exploring Stakeholder Value Models Via Interactive Visualization," 12th Conference on Systems Engineering Research, Redondo Beach, CA, March 2014.
- Richards, M.G., "Multi-Attribute Tradespace Exploration for Survivability," Doctor of Philosophy Dissertation, Engineering Systems Division, MIT, June 2009.
- Ross, A.M., Fitzgerald, M.E., and Rhodes, D.H., "Interactive Evaluative Model Trading for Resilient Systems Decisions," 14th Conference on Systems Engineering Research, Huntsville, AL, March 2016.
- Ross A.M., Rhodes D.H., and Fitzgerald M.E., "Interactive Value Model Trading for Resilient Systems Decisions," 13th Conf on Sys Eng Research. (CSER15). Hoboken, NJ, March 2015.
- Ross, A.M., McManus, H.L., Rhodes, D.H., Hastings D.E. "A role for interactive tradespace exploration in multi-stakeholder negotiations", AIAA Space 2010. Anaheim, CA, September 2010.
- Ross, A.M., McManus, H.L., Rhodes, D.H., Hastings, D.E., and Long, A.M., "Responsive Systems Comparison Method: Dynamic Insights into Designing a Satellite Radar System," AIAA Space 2009, Pasadena, CA, Sep 2009.
- Ross A.M., Rhodes D.H.. Using natural value-centric time scales for conceptualizing system timelines through epoch-era analysis. INCOSE Int'l Symp 2008, Utrecht, the Netherlands, June 2008.
- Ross, A.M., McManus, H.L., Long, A., Richards, M.G., Rhodes, D.H., and Hastings, D.E., "Responsive Systems Comparison Method: Case Study in Assessing Future Designs in the Presence of Change," AIAA Space 2008, San Diego, CA, September 2008.
- Ross A.M., Managing unarticulated value: changeability in multi-attribute tradespace exploration. PhD in Engineering Systems. Cambridge, MA: MIT, 2006.
- Rouse, W.B., Modeling and Visualization of Complex Systems and Enterprises, Exploration of Physical, Human, Economic and Social Phenomena, 2015.
- Sarter, Nadine B, and David D Woods. "Mode Error in Supervisory Control of Automated Systems." Proceedings of the Human Factors and Ergonomics Society Annual Meeting. Santa Monica: Human Factors and Ergonomics Society, 1992. 26-29.
- Sarter, Nadine B, David D Woods, and Charles E Billings. "Automation Surprises." In Handbook of Human Factors & Ergonomics, 2nd Ed, 1926-1943. Hoboken: John Wiley & Sons, 1997.
- Schaffner MA. Designing systems for many possible futures: the RSC-based method for affordable concept selection (RMACS), with multi-era analysis. SM in Aeronautics and Astronautics. Cambridge, MA: MIT, 2014.
- Simon, H., "The architecture of complexity", Proceedings of the American Philosophical Society, 106, 6, 467– 482, 1962.
- Skitka, Linda J, Kathleen L Mosier, and Mark D Burdick. "Accountability and Automation Bias." International Journal of Human-Computer Studies, 2000: 701-717.
- Smaling, R., "Fuzzy Pareto Frontiers in Multidisciplinary System Architecture Analysis", 10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, Albany, NY, September 2004.
- Spero E, Avera MP, Valdez PE, Goerger SR. "Tradespace exploration for the engineering of resilient systems", 12th Conf on Sys Eng Research. (CSER14). Redondo Beach, CA, March 2014.
- Standards Development Organization and Advancement Act of 2004, U. S. Congress, 2004.
- Stim, R., and Stephan, F., What To Do If An NDA Is Violated, NDAS For Free, Available: [www.ndasforfree.com/4StepstoTake.html](http://www.ndasforfree.com/4StepstoTake.html) [Accessed: 2/2/16].

- Strauch, Barry. "Projected Flight Path Displays and Controlled Flight Into Terrain Accidents." Digital Avionics Systems Conference. Bellevue, WA: IEEE, 1998. E43/1 - E43/8.
- Suh, N. P., "Complexity—theory and applications", Oxford University Press, Oxford, UK, 2005.
- Thomas, J. and Cook, K. "Illuminating the Path: The R&D Agenda for Visual Analytics National Visualization and Analytics Center," National Visualization and Analytics Center, 2005.
- Thomas, J., Director, USDHS National Visualization and Analytics Center, "Visual Analytics: An Agenda in Response to DHS Mission Needs," 2007.
- Thomson, K., de Chernatony, L. Arganbright, L. and Khan, S., "The Buy-in Benchmark: How Staff Understanding and Commitment Impact Brand and Business Performance". Journal of Marketing Management 1999; vol. 15, 819-835.
- Viscito, L., and Ross, A.M., "Quantifying Flexibility in Tradespace Exploration: Value-Weighted Filtered Outdegree," AIAA Space 2009, Pasadena, CA, September 2009.
- Warwick. USAF Selects Lead Programs for 'Digital Twin' Initiative. Aviation Week & Space Technology; 2015
- West, T.D. and Pyster, A. Untangling the Digital Thread: The Challenge and Promise of Model-Based Engineering in Defense Acquisition. INCOSE INSIGHT 2015; vol. 18, 45-55.
- Wickens, Christopher D, and Amy L Alexander. "Attentional Tunneling and Task management in Synthetic Vision Displays." The International Journal of Aviation Psychology, 2009: 182-199.
- Wiener, Earl L. Human Factors of Advanced Technology ("Glass Cockpit") Transport Aircraft. Moffett Field: National Aeronautics and Space Administration, 1989.
- Wodzicki, K., Schwämmlein, E., Cress, U. and Kimmerle, J. "Does the Type of Anonymity Matter? The Impact of Visualization on Information Sharing in Online Groups", Cyberpsychology, Behavior, and Social Networking 2010; vol. 14, 157-160.
- Wong, P.C. and Thomas, J. "Visual Analytics". IEEE Computer Graphics and Applications, Vol. 24, Issue 5, 2004 Page(s): 20–21.
- Wright, Stephen, and David O'Hare. "Can a glass cockpit display help (or hinder) performance of novices in simulated flight training?" Applied Ergonomics 47 (2015): 292-299.
- Zacharias, G.L., MacMillan, J., Van Hemel, S. B., Behavioral Modeling and Simulation, 2008.
- Zimmerman, P., A Framework for Developing a Digital System Model Taxonomy. 18th Annual NDIA Systems Engineering Conference; 2015.
- Zimmerman, P., MBSE in the Department of Defense. Office of the Deputy Assistant Secretary of Defense for Systems Engineering, DoD, 2015.
- Zimmerman, P., Modularity and Open Systems: Meaningful Distinctions. 18th Annual NDIA Systems Engineering Conference; 2015.
- Zulkepli, J., Eldabi, T., Mustafee, N., Hybrid simulation for modelling large systems: an example of integrated care model, Proceedings of the 2012 Winter Simulation Conference, 2012.