DESIGN OF FLEXIBLE TECHNOLOGY REFRESH PLANS FOR MILITARY OPEN SYSTEM ARCHITECTURES

A Thesis
Presented to
The Academic Faculty

by

Eric M. Zellers

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Aerospace Engineering

Georgia Institute of Technology
May 2016

Copyright © 2016 by Eric M. Zellers

المنارة للاستشارات

www.manaraa.com

DESIGN OF FLEXIBLE TECHNOLOGY REFRESH PLANS FOR MILITARY OPEN SYSTEM ARCHITECTURES

Approved by:

Dimitri Mavris, Professor, Advisor
School of Aerospace Engineering
Georgia Institute of Technology

Daniel Schrage, Professor
School of Aerospace Engineering
Georgia Institute of Technology

Charles Domercant, Ph.D.
Aerospace Systems Design Laboratory
Georgia Institute of Technology

John Salmon, Assistant Professor
School of Mechanical Engineering
Brigham Young University

Mr. Gary O'Neill
Georgia Tech. Research Institute

Date Approved: 15 March, 2016

www.manaraa.com

To my father,

John Roy Zellers

www.manaraa.com

ACKNOWLEDGEMENTS

The development of this thesis has been both rewarding and challenging. This effort would not have been possible without the constant aid and support of my friends and family. First and foremost, I would like to thank my father for the perpetual love, faith, and confidence he's shown me. Passing this milestone in life is truly a momentous occasion, but I will always feel a deep sense of remorse knowing that I will not be able to share this moment with him. He is, and will continue to be, deeply missed. I would also like to thank my mother for her love and compassion in these trying times, and for always trying to find a way to ease my burdens. To my brother, thank you for always finding the time and patience to lend a hand with the many technical hurdles that arose along the way. Special thanks to my Uncles, Jim Zellers and Jim Devlin, and my cousin Erin Devlin, who took it upon themselves to look after me throughout this process.

I am indebted to my many colleagues for their valuable feedback, much needed criticism, and positive support. This includes all of my friends and colleagues at ASDL, both past and present. There are simply too many important people in this group to give everyone the credit they deserve, but I would like to especially thank Nathan Knisely, Alicia Sudol, Patrick Dees, Eugina Ramos, Daniel Cooksey, and Nicole Arockiam for their willingness to serve as an essential sounding board for my ideas. Your friendship and endless patience have been invaluable in this process.

I would also like to express gratitude to my committee; Professors John Salmon and Daniel Schrage, Dr. Charles Domercant, and Mr. Gary O'Neill. Your excellent feedback and advice were essential to keeping me on track throughout this process. I owe a special debt of gratitude to Dr. Charles Domercant, whose council was critical

المنارة للاستشارات
www.manaraa.com

to moving this work through the early phases of development. Last, and certainly not least, I am tremendously grateful to my advisor and mentor, Dr. Dimitri Mavris. Over the past five years, I have come to admire both the depth of his knowledge, and his passion for developing the next generation of engineering. I could not have asked for a more supportive or dedicated advisor, and I am eternally grateful the many ways in which he shaped this experience.

www.manaraa.com

TABLE OF CONTENTS

المنارة للاستشارات
www.manaraa.com

المنارة للاستشارات
www.manaraa.com

المنارة للاستشارات
www.manaraa.com

المنارة للاستشارات
www.manaraa.com

LIST OF TABLES

المنارة للاستشارات

www.manaraa.com

LIST OF FIGURES

www.manaraa.com

المنارة للاستشارات
www.manaraa.com

المنارة للاستشارات
www.manaraa.com

المنارة للاستشارات
www.manaraa.com

المنارة للاستشارات
www.manaraa.com

SUMMARY

Military acquisition programs have long been criticized for the exponential growth in development costs required to generate modest improvements in capability. One of the most promising reform efforts to address this trend is the open system architecture initiative, which uses modular design principles and commercial interface standards as a means to reduce the cost and complexity of upgrading systems over time. While conceptually simple, this effort has proven to be exceptionally difficult to implement in practice. This difficulty stems, in large part, from the fact that open systems trade additional cost and risk in the early phases of development for the option to infuse technology at a later date, but the benefits provided by this option are inherently uncertain. Practical implementation therefore requires a decision support framework to determine when these uncertain, future benefits are worth the cost and risk assumed in the present, but there is ample evidence to suggest that existing design methods are insufficient to address this need.

The objective of this research is to develop a Military Acquisition INspired FRamework for Architecture Modeling and Evaluation that resolves this gap by providing an approach to measure the expected costs, benefits and risks associated with open systems. This work is predicated on three assumptions: (1) the purpose of future technology infusions is to keep pace with the uncertain evolution of operational requirements, (2) successful designs must justify how future upgrades will be used to satisfy these requirements, and (3) program managers retain the flexibility to adapt prior decisions as new information is made available over time. With that in mind, this methodology proposes a new technique for codifying operational requirements as a capability road map, as opposed to the "worst case" scalar values used in classical design methods.

المنارة للاستشارات
www.manaraa.com

A novel adaptation of existing technology forecasting techniques is then proposed as a means to determine how future technological improvements could be used to efficiently satisfy the needs expressed in this road map, and a new performance measure is proposed to quantify the relative value of alternative refresh strategies. Finally, a series of decision support heuristics inspired by methods in the field Real Options are integrated with an automated search procedure to identify strategies that facilitate flexible decision making as a hedge against uncertainty.

The proposed methodology is then applied to an example scenario for an aerial Intelligence, Surveillance, and Reconnaissance platform with the potential to upgrade its sensor suite in future increments. The capability road map for this scenario is adapted from real world trade studies performed by the Department of Defense's Information Dominance team, and the forecasting model is developed by evaluating technological progression in commercial image processing technology over the last decade. Specific questions addressed in this study are how the timing and selection of future technology infusions should be structured to best satisfy alternative preferences for cost, performance, and risk. In addition, the study demonstrates that the relative advantages and drawbacks, in terms of the performance metrics developed in this work, between open and integrated system architectures can be presented in the context of a cost-effectiveness framework that is currently used by acquisition professionals to manage complex design decisions. This experiment concludes with the observation that the proposed methodology can objectively identify and aggregate the myriad of factors impacting an arbitrary open system design problem into a single, intuitive visualization. As this capability is lacking in existing methods, it lends considerable support to the thesis that the proposed methodology is a superior approach.

المنارة للاستشارات
www.manaraa.com

CHAPTER I

MOTIVATION

[The] DoD has been forced to cancel one unaffordable program after another to live within budget constraints. When taken as a whole, it is obvious that continuing "business as usual" in defense acquisition is not sustainable...our buying strategies must adapt to this new reality and recognize that the costs of our weapon systems must assume a more prominent role in the decision process; our nation's future depends on it [96]

Frank Kendall, Under Secretary of Defense - Acquisition, Technology, and Logistics

The first obligation of any sovereign nation is to provide for the safety of its citizens, and an effective military is the primary instrument to fulfill this obligation. Yet, security is one of many challenges, much like health care or education, a nation must satisfy from a finite pool of resources. Maintaining an effective military therefore requires, among other things, ensuring that present and future forces are equipped with the proper systems at an appropriate cost. The notion of a cost-effective military is an integral part of the United States' National Security Strategy (NSS), but efforts to control the cost of developing military systems has proven to be an exceptionally difficult challenge. As noted by Frank Kendall above, the status quo in defense acquisition cannot be maintained indefinitely; new engineering and management methods will be necessary to satisfy future capability requirements with constrained resources [96]. This research is intended to be a contribution toward the development of these new methods.

1.1 Cost Growth in Military Systems

Technological superiority has been, and will continue to be, a cornerstone of the American defense policy. Objectively, this policy appears to be well founded. At the strategic level, technological superiority enabled a numerically smaller NATO force

www.manaraa.com

to deter Soviet aggression in Western Europe during the Cold War, and is credited with deterring open hostility from other major powers today. The overwhelming success of American forces in the first Gulf War also validated the policy's utility at the tactical and operational levels of warfare. Yet, the costs associated with an incremental increase in technological sophistication have grown at an exponential rate. For example, consider the long term trend in aircraft cost shown in Figure 1. The solid line indicates a real annual growth rate of approximately 5.5% [1], which corresponds to unit costs nearly doubling every ten years [113].

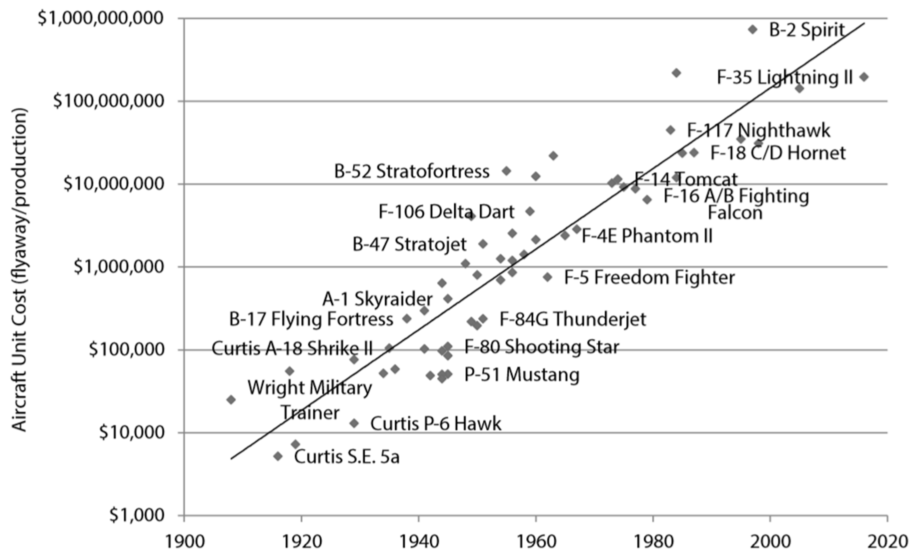


Figure 1: Rapidly Increasing Cost of Aircraft [100]

Over the same time period, however, the American economy has grown at an average annual rate of approximately 2.9%[2]. If one assumes that the defense budget grows at the same rate as the greater economy, such that a 2.9% increase in Gross Domestic Product (GDP) lead to a 2.9% growth in defense spending[3], then the severity of the problem becomes clear. Costs are increasing at nearly twice the

[1] The trend in the figure demonstrates a nominal growth rate of 10%, which is reduced to 5.5% after accounting for inflation over the same time period.

[2] Inflation data retrieved from http://www.multpl.com/us-real-gdp-growth-rate/table/by-year

[3] This is also a highly conservative assumption, as the share of U.S. GDP allocated to defense steadily declined from 15% in 1952 to 3.8% in 2013.

المنارة للاستشارات
www.manaraa.com

rate at which new resources can be allocate to pay for those costs, and this difference has been compounding for decades. Moreover, Figure 2 shows that the response to this imbalance has, quite predictably, been a decrease in the number of units acquired over time. Norman Augustine, former Chief Executive of Lockheed Martin and acquisitions luminary, famously characterized the implications of this trend [10].

> In the year 2054, the entire defense budget will purchase just one aircraft. The aircraft will have to be shared by the Air Force and Navy 3 $\frac{1}{2}$ days per week except for leap year, when it will be made available to the Marines for the extra day

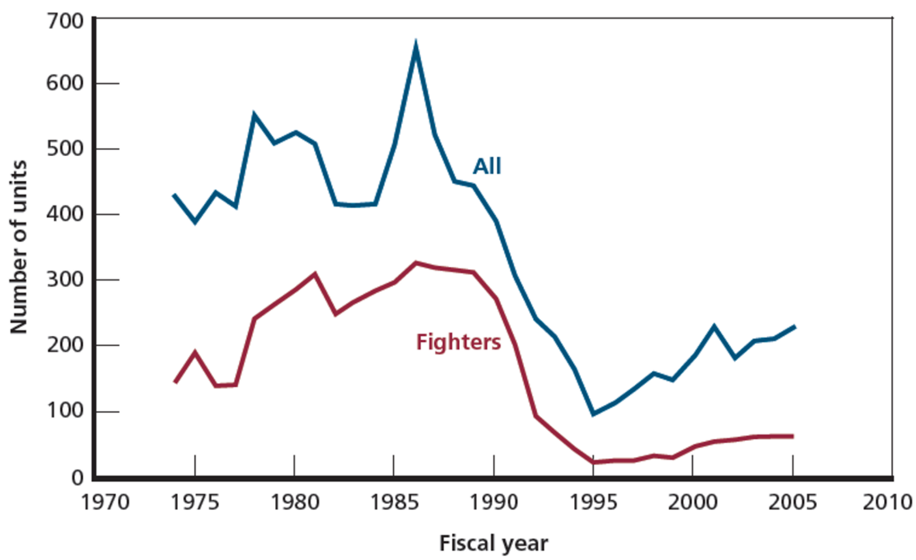


Figure 2: Change In Aircraft Production Over Time [113]

1.2 Previous Reform Efforts

The trend of exponential growth in unit acquisition costs is not limited to aircraft, nor is this trend a recent phenomenon [9, 23]. Scholars, law makers, and military professionals have studied the perennial challenge of cost escalation in defense acquisition for over 50 years. From 1960 to 2009, more than 27 major studies of the defense

المنارة للاستشارات
www.manaraa.com

acquisition system were commissioned by various stakeholders, as well as numerous, yet no less noteworthy, studies from the Government Accountability Office (GAO) and the Congressional Research Service (CRS) [64, 66]. Collectively, these efforts have reformed nearly every facet of the acquisition system in some meaningful way. Though a complete review of these efforts is beyond the scope of this analysis, there are three important management programs that have made significant progress in reducing cost growth.

- Cost as an Independent Variable (CAIV)
- Earned Value Management (EVM)
- Nunn-McCurdy Legislation

There is a common saying in program management literature - *fast, cheap, or good; pick two, but the third will be what it must to accommodate the other two.* CAIV embraces this argument to form the first layer of defense by fixing cost early in the design phase, and performing trade studies to determine the possible alternatives for schedule and performance under that constraint [67, 64]. Once a program is under way, EVM tracks the progress of the program in real time. This is accomplished by decomposing the development program into a set of elementary tasks, which are paired with estimates of the time and cost required to complete them. Substantial deviations from the initial estimate are reflected in various EVM metrics, which allow Program Managers (PM) to address issues before they compound to a point of critical mass [84, 156]. Finally, an external team reviews development programs every quarter under the Nunn-McCurdy Act of 1983 [4]. If these reviews find that cost projections

[4]410 U.S.C. 2433. The statutory provision is known as Nunn-McCurdy because it was first introduced by Senator Nunn and passed as a 1-year provision as part of the Department of Defense Authorization Act, 1982. 127 Cong. Rec. 9760-63 (1981); Pub. L. No. 97-86, 917. The following year, Representative McCurdy introduced a permanent provision based on Senator Nunn's provision, which was enacted as part of the Department of Defense Authorization Act, 1983. 128 Cong. Rec. 18345-48 (1982); Pub. L. No. 97-252, 1107 [67].

المنارة للاستشارات
www.manaraa.com

have grown to be greater than 15% of the baseline estimate, then the PM is obligated to submit a report to Congress detailing the new cost estimate, what caused the increase, and how that cause will be corrected. Thus, these initiatives provide a layered defense against cost growth before and during development, and well as with internal and external oversight.

1.3 Enduring Challenges

Despite this elaborate defense against cost growth, the 2013 GAO survey of defense programs found that 56% of programs still experienced cost growth greater than 15% of their respective baseline estimates [68]. Though this performance is better than the historical average, it is clear that challenges persist. Unfortunately, the reasons for this persistence are quite unclear. Among the competing explanations, however, are three factors that warrant further consideration due to their relationship with existing Systems Engineering (SE) methods: gold plating, vendor lock, and integration of Commercial-Off-the-Shelf Components (COTS) [19].

Gold plating, also known as "requirements creep", is the phenomenon where military services specify demanding performance requirements, which in turn leads to the addition of design features whose cost exceeds their expected value [64]. This is believed to be caused by the fact that, due to the cost and complexity of development, advanced military systems have extremely long lifecycles - often in excess of 40 years[5]. In order to ensure that the system is operationally useful throughout its planned service life, design requirements are often derived from what is expected of the system at the end of its lifecycle. Forecasting this far into the future, however, introduces significant uncertainty. Establishing a high level of confidence that a design will satisfy its terminal requirements, given their high degree of uncertainty, leads to the demanding requirements at the heart of the gold-plating phenomenon.

[5]For example, the first Nimitz-class aircraft carrier, USS Nimitz (CVN68), was commissioned in 1975 and is still in service as of 2015. - http://www.nimitz.navy.mil

المنارة للاستشارات
www.manaraa.com

Dr. Paul Kaminski, former Under Secretary of Defense for Acquisition and Logistics Technology (USD AT&L), summarized this as follows [19]:

> Gold plating is bad, especially when driven by a very long acquisition cycle which creates the incentive to include everything but the kitchen sink in the requirements for a new system because who knows when we will have the next opportunity for enhancement or replacement.

Vendor lock describes a situation in which an organization becomes dependent on a single manufacturer or supplier for a given product, and cannot move to another vendor without incurring substantial costs or inconvenience. These situations are analogous to allowing the vendor to establish a monopoly in the marketplace, which they can then leverage to establish noncompetitive prices. Instances of vendor lock commonly occur when the vendor retains proprietary data and Intellectual Property (IP) rights over aspects of the design. Consequently, the government effectively has two options when changes in requirements force the system to be modified - accept the prices offered by the vendor, or purchase the IP or proprietary information to facilitate a competitive bidding process. Maintaining effective competition therefore requires the government to maintain some measure of control over the technical data rights embedded in the systems they acquire [55, 176]. The DoD handbook for acquiring data rights provides a very direct statement of the significance of this concept [136]:

> If we do not acquire sufficient rights in technical data and computer software prior to award, we may relinquish the opportunity to enhance competition....thereby locking ourselves into a position whereby the incumbent can force us to pay an exorbitant price years or decades hence

Finally, military development programs often require components designed to serve a military specific function under extreme operating conditions. These components do not exist in commercial markets, and must therefore be custom developed

www.manaraa.com

for the military. The benefit of these products is that military can specify exactly what it wants; the drawback is that it must also bear the full burden of fixed costs for design and manufacturing set-up in addition to the per unit cost of the final product. These costs can be substantial. For example, the coffee maker on the C-5 transport was low pressure certified and rated to 50 G's acceleration, but cost $7,622 per unit. Purchasing COTS components, on the other hand, allows fixed costs to be amortized across a larger base. Moreover, COTS components tend to coalesce around commonly accepted standards, which makes them easier to change out if requirements or preferences change over time [28, 51, 99]. The desire to capitalize on the virtues of COTS components wherever possible is a major emphasis in the acquisition community, and is currently mandated in DoD acquisition regulations [52]:

> DoD Components shall seek the most cost-effective solution over the system's life cycle. They shall conduct market research and analysis to determine the availability...of commercially available products, services, and technologies, from domestic or international sources, or the development of dual-use technologies

1.4 Proposed Solution

Each of the cost drivers discussed above are significant challenges in present day acquisition programs [19]. This is due, to great extent, because they are technical challenges pertaining to SE methods as opposed to managerial challenges considered by CAIV, EVM, and Nunn-McCurdy oversight. A technical approach is therefore required to address these issues. The acquisitions community has proposed such a technical solution in the form of the Open System Architecture (OSA) initiative.

The OSA initiative was mandated as a best practice, one that must be executed where possible, in a 2003 revision to DoD Regulation 5000.2-R [64]. These regulations have been revised periodically to clarify and strengthen the mandate, but the basic

www.manaraa.com

concept remains unchanged from its 2003 inception. In essence, Open Architecture (OA) systems leverage the basic design concepts of physical modularity and functional partitioning of system capabilities in order to allow systems to be easily modified over time. In addition, the OSA initiative mandates that the interfaces connecting modules to the greater system architecture conform, where possible, to existing commercial standards, and that these standards be made widely available to industry. Thus, when a component or sub-system requires some degree of modification, it can be replaced quickly and affordably with either a COTS component, or through a competitive bidding process [13, 14, 12, 70, 75].

A commonly cited example of an open interface are the tires on a car [75]. The mechanical interface connecting the tire to the vehicle is governed by well established commercial standards, which are typically listed on the tire's sidewall under a common convention. Whenever a car needs new tires, the owner does not need to return to the original manufacturer to procure the original tire. Rather, the owner can use the information on the tire's sidewall to identify numerous alternatives from multiple vendors, confident that each option can be integrated with equal effort. Strong competition between these vendors keeps prices low and quality high, thereby making it highly likely that the tires available in the present are Pareto efficient compared to the original set. Moreover, if the owner finds that their preferences have changed, they have the opportunity to gain additional value by selecting an option that more closely matches their current needs.

The argument for OSA design becomes more compelling when considering the impact such a system would have on the concepts of gold plating, vendor lock, and the need for custom components. Recall that gold plating is especially problematic when the system is difficult to modify, as requirements must be derived from the end of the system's life life to ensure it remains useful over its design life. By making the system easier to modify, OSA allows requirements to be derived from an earlier

المنارة للاستشارات
www.manaraa.com

point in the system's life, where uncertainty is dramatically reduced. Further, because the interface standards are widely distributed, all commercial vendors are able to bid for modification contracts. This relegates the impact of vendor lock to the tightly integrated portions of the architecture that are unlikely to require substantial modification over time. Finally, the mandate that interface standards conform to commercial standards allows the DoD to maximize integration of COTS components and minimize the need for custom designs.

Though conceptually simple, the OSA initiative has proven to be quite difficult to execute in practice. A recent review of OSA design attempts in military acquisition programs revealed the following barriers to success [60]:

- No consistent definitions on what it means to be "open", or what the appropriate level of "openness" is in a system
- There is a financial cost for imposing OA constraints on a portion of a given design, to include the costs of purchasing proprietary information and configuration management.
- OA implementation typically does not have its own line item in the program budget. This creates an opportunity cost where program funds must be diverted from other worthy activities to pay of OA development
- The SE methods supporting OA design are still immature, which increases the development risk of OA design; this risk is difficult to quantify
- The benefits of OSA are longer term and are subject to considerable uncertainty
- There is an inherent conflict for the PM. PM's are typically judged on short term results (*e.g. EVM metrics*), but the benefits of an OA may be years in the future

للاستشارات المنارة
www.manaraa.com

- Short term costs and risk of OA development can be mitigated by not emphasizing OA, as the end product provides the same functionality to the war fighter
- Prime contractors have a conflict of interest; they can be tasked with identifying which system elements should be open, but vendor lock is in their financial interest

In summation, OSA has great potential to resolve long standing challenges that contribute to the greater problem of cost growth in military systems. However, new methods are required to determine if the uncertain benefits provided in the future are worth the cost and risk today. These methods must also be standardized and integrated into the existing SE structure in order to achieve greater acceptance within the acquisition community. This research effort will therefore focus on developing such a method.

To that end, the remainder of this work will be structured as follows. Chapter Two will present a review of the Systems Engineering (SE) methods currently used by the acquisition community to manage the development of both closed and open system architectures. These methods will then be juxtaposed with the barriers to success denoted above to refine the problem into a formal research objective. Chapter Three will then identify a series of observations from Chapter Two related to the research objective. These observations will help refine the objective into a set of research questions that must necessarily be addressed in order for this work to be successful. Chapters Four, Five, and Six will then consider each of these research questions in series to determine if sufficient methods exist within the academic literature to resolve the gap in military acquisition methods. If academic methods fail to provide an appropriate solution to a research question, then a novel approach will be created to address the gap. Chapter Seven will then integrate the methods used to resolve the

المنارة للاستشارات
www.manaraa.com

research questions into an overall methodology that is believed to satisfy the overall objective. Chapter Eight will then apply the methodology to a real world problem in order to validate this thesis. Finally, Chapter Nine will review the contributions made through the development of this methodology and elaborate on the potential for future work.

المنارة للاستشارات
www.manaraa.com

CHAPTER II

ACQUISITION PRACTICES

An appropriate starting point for the development of a methodology to manage the complexities of OSA design is to consider the SE methods currently used by acquisition professionals. In broad terms, these existing methods follow a layered approach. The first layer is the higher level acquisition system, the so-called "Big A", which manages the generation of requirements, budgeting, and the engineering development process. The next layer is a generic Analysis of Alternatives (AoA) methodology applicable to all military systems, regardless of whether the architecture is open or closed. As the name implies, this phase identifies a useful set of alternatives to meet the capability requirements for which the system under development is intended to satisfy, and ends with the selection of a particular system concept. The final layer activates once a concept is selected for development. The intent of these methods is two-fold: (1) ensure that the development process is structured in such a way that it is possible to identify system elements that would benefit from future upgrades, and (2) determine whether a given element should be modularized or integrated into the greater system architecture. With that in mind, the remainder of this chapter will review how these methods are structured, and then compare them to the barriers to success identified in Chapter One. This comparison will provide insight on the gaps that must be addressed by the methodology under development.

2.1 Defense Acquisition Overview

The defense acquisition system is the management process through which the DoD provides effective, affordable, and timely systems to the military [52]. These processes provide the overarching structure which identifies the military capabilities required

المنارة للاستشارات
www.manaraa.com

to achieve high level NSS objectives, and develops systems to provide those capabilities. When a materiel[1] solution is required, these processes cover every facet of the system life cycle, to include: concept generation, resource allocation, engineering development, test and evaluation, production, deployment, sustainment, and disposal. In order to manage the diverse needs of such a far-reaching system, management activities are allocated to three interrelated decision support systems.

- Planning, Programming, Budgeting, and Execution (PPBE)
- Joint Capabilities Integration Development System (JCIDS)
- Defense Acquisition System

The interacting nature of these decision support systems is depicted in Figure 3. In addition, Figure 3 highlights an important distinction in nomenclature. The "Big A" process is the highest level conception of the acquisition process, which encompasses all three decision support systems. The defense acquisition system, the so called "Little A", on the other hand, is a specific decision support system. DoD regulation 5000.01 clarifies this distinction with a description of the Big-A process [52]:

> The Defense Acquisition System exists to manage the nation's investments in technologies, programs, and product support necessary to achieve the National Security Strategy and support the United States Armed Forces Forces. The investment strategy of the Department of Defense shall be postured to support not only today's force, but also the next force, and future forces beyond that. The primary objective of defense acquisition is to acquire quality products that satisfy user end needs with measurable

[1] *Materiel* is a generic word for equipment. It is inherently plural and is distinguished from material, which is the physical product from which systems are composed. For example, military aircraft are materiel; whereas the aluminum, steel, titanium, etc. which compose the aircraft are materials [25].

المنارة للاستشارات
www.manaraa.com

improvements to mission capability and operational support, in a timely manner, and at a fair and reasonable price.

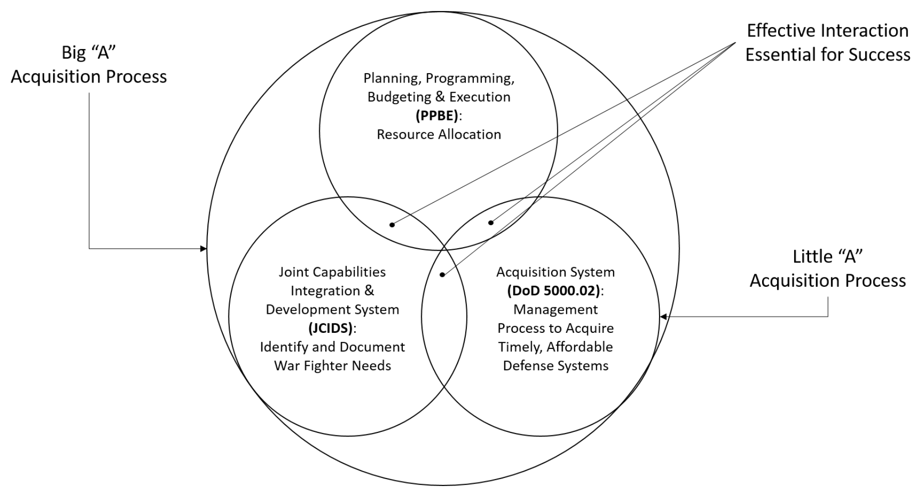


Figure 3: Defense Acquisition Decision Support Systems [143, 59]

The Little-A acquisition process provides the management structure through which system concepts are brought to fruition. Some of the objectives that govern this process are [59]:

- Flexibility
- Responsiveness
- Innovation
- Discipline
- Streamlined and Effective Management
- Cost and Affordability
- Cost Realism
- Integrated Test and Evaluation

www.manaraa.com

For management purposes, all DoD development programs are grouped according to the expected sunk cost required for Research, Development, Testing, & Evaluation (RDT&E) and overall procurement (RDT&E and production). These categories are referred to simply as Acquisition Categories (ACAT), and the rule set for determining which ACAT a program belongs to is provided in Figure 4. Of note in this figure is the fact that any program with sufficient risk may be elevated to a higher ACAT, even though the fiscal threshold is not met. Higher ACAT levels possess a higher level of statuary oversight, in the form of the Milestone Decision Authority (MDA), which is intended to help mitigate excessive risk. Given the breadth of information and management challenges implicit in the development of any complex system, however, it must be recognized that this oversight must be judicious. As such, there is some criticism that the current ACAT paradigm does not go far enough in addressing development risks [131].

Category	Criteria for Designation	Decision Authority
ACAT I	• Major Defense Acquisition Programs - RDT&E total expenditure of more that $365M, or - Procurement total expenditure of more than $2.190B • MDA designation as special interest	• ACAT ID: USD(AT&L) - Reviewed by the Defense Acquisition Board (DAB) • ACAT IC: Component head, or Component Acquisition Executive (CAE) (cannot be further delegated) - Reviewed by component HQ
ACAT II	• Does not meet ACAT I criteria • Major System - RDT&E total expenditure of more than $140M, or - Procurement total expenditure of more than $660M • MDA designation	• CAE or the individual designated by the CAE • Reviewed in accordance with component policy
ACAT III	• Does not meet ACAT II or above criteria	• Designated by the CAE at the lowest appropriate level • Reviewed in accordance with component policy

Figure 4: Aquisition Categories for Weapon Systems (FY 2000 Dollars) [59]

المنارة للاستشارات
www.manaraa.com

Regardless of the ACAT, and by extension the MDA, of a particular program, DoD policy stipulates that a PM be assigned to oversee development [52]. Program management represents the synthesis and integration of a myriad of functional disciplines, including business and financial management, logistics, Systems Engineering, software management, test and evaluation, manufacturing, etc. These disciplines must work together in order to fully realize the overall goals of the development program [143, 59].

2.1.1 JCIDS: Joint Capabilities Integration Development System

For much of its history, the defense acquisition system was structured around the belief that weapons systems should be designed for a individual branch of the military to counter a specific threat. This service-centric, stove-piped approach often led to systems that lacked interoperability, were duplicative, or did not fill critical gaps. In 2001, the DoD came to the conclusion that the threat-based model should be replaced by a capabilities-based model focusing on how an adversary will fight, rather than who the adversary might be or where a war might be fought. Change came two years later, when the legacy requirement generation system was replaced with the Joint Capabilities Integration and Development System (JCIDS) process [164].

JCIDS is a collaborative process overseen by the Joint Requirements Oversight Council (JROC[2]), whose purpose is to ensure that any capabilities required by the warfighter to successfully execute their mission are identified, along with their associated performance criteria [25]. Figure 5 shows the role of JCIDS in the requirements generation process. The process starts with an enumeration of the missions, known as Concepts of Operation, the DoD could be called on to perform through its role in the NSS. These missions are conceptually reduced to a set of capabilities, and

[2]JROC is chaired by the Vice Chairman of the Joint Chiefs of Staff. Members include the vice chiefs of staff of the Army and Air Force, the Vice Chief of Naval Operations, and the Assistant Commandant of the Marine Corps.

المنارة للاستشارات
www.manaraa.com

any perceived gap in the required capabilities, present or future, will require either a corrective action, or a deliberate assumption of risk. Where corrective actions are required, JCIDS is the mechanism employed to identify an optimal solution through any combination of changes to Doctrine, Organization, Training, Materiel, Leadership and Education, Personnel and Facilities (DOTMLPF). Once a solution, or set of candidate solutions, are identified, JCIDS will determine the relevant performance metrics and generate a series of documents used to inform subsequent phases of development. These documents are [164, 25, 32, 33]:

- *Initial Capabilities Document (ICD):* Provides the definition of the capability, and a description of how it fits into the broader operational mission concept. The ICD is also used to support the materiel and technology development phases, as well as the Milestone A review.

- *Capability Development Document (CDD):* Supports the Milestone B review by providing more detail on the materiel solution previously described in the ICD. The CDD also provides objectives and targets for system attributes against which the delivered capability will be measured.

- *Capability Production Document (CPD):* Provides updated performance requirements for system attributes based on lessons learned during the engineering and manufacturing phases of system development. The CPD is used to inform Milestone C decisions before a program enter Low-Rate Initial Production (LRIP).

2.1.2 PPBE: Planning, Programming, Budgeting, and Execution

The DoD has a finite pool of resources to satisfy an almost unlimited demand from the various military services trying to fulfill their respective roles in the NSS. The PPBE decision support systems provide the resource allocation process to determine how

المنارة للاستشارات
www.manaraa.com

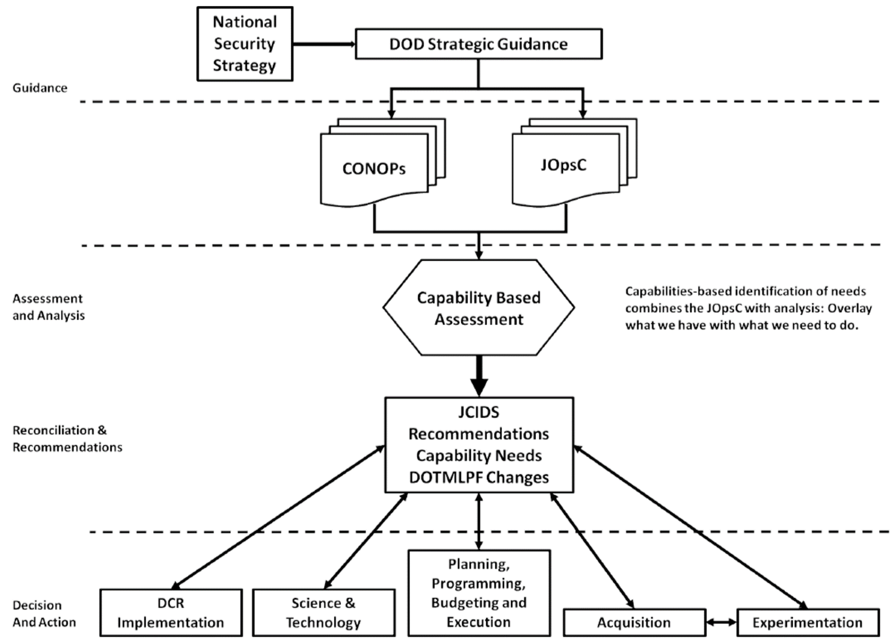


Figure 5: Joint Capabilities Integration and Development System [59]

resources are to be distributed across these competing needs. This process, shown in Figure 6, functions in four phases [25]:

- Planning, Programming, Budgeting, and Execution (PPBE) Process
- Enactment
- Apportionment
- Allocation / Execution

The purpose of the PPBE is to produce the DoD portion of the President's national budget. Enactment and Apportionment fall under the purview of the Congressional budgeting process, and are therefore beyond the control of the acquisition system. The final phase, Allocation and Execution, mark the start point of the system development process.

المنارة للاستشارات
www.manaraa.com

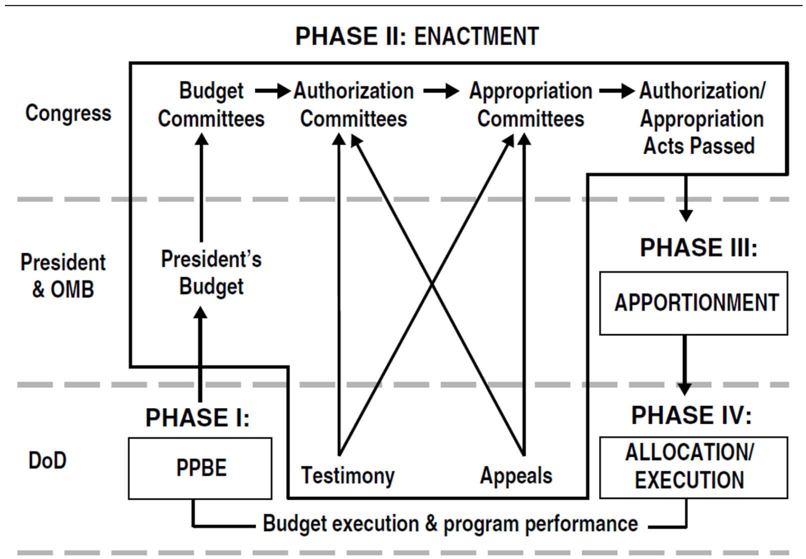


Figure 6: Resource Allocation Process

The PPBE cycle begins with a collaborative planning effort by the Secretary of Defense, the Joint Staff, and DoD components to develop a resources-informed articulation of the capabilities required to achieve the major NSS objectives. The result is encapsulated in a formal defense planning guidance document used to lead the overall planning process. Publication of this planning guidance marks the start of both the the programming and budgeting phase. In the programming phase, each DoD component attempts to develop a balanced set of programs to respond to the priorities provided in the defense planning guidance. This internal process culminates in a Program Objective Memorandum (POM), which provides a detailed description of the programs proposed and a five year plan for the time-phased allocation of resources intended for each program. Budgeting occurs in parallel with the Programming phase. During this phase, each component develops a Budget Estimate Submission (BES) to convert programmatic decisions, along with supporting documentation, into the congressional appropriation format. This document is submitted alongside the POM. The final phase, execution, is a review process that provides feedback to senior leadership on the effectiveness of prior resource allocation decisions, and is supported by

المنارة للاستشارات
www.manaraa.com

various metrics used to measure actual output versus planned performance [25].

2.1.3 Acquisition Life Cycle

The life cycle of a system from its initial conception to Operations & Support (O&S) is separated into different phases by decision points known as *milestones*. Three milestones must be successfully navigated before a system reaches its Initial Operating Capability (IOC). These milestones, designated A, B, and C respectively, are depicted in Figure 7 [51, 59].

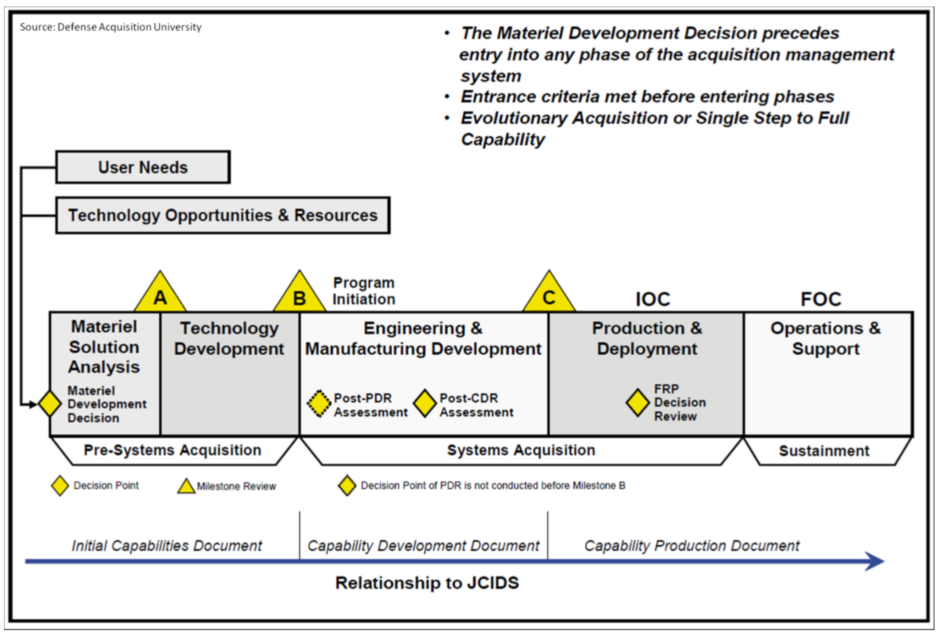


Figure 7: Acquisition Lifecycle [59]

The designated MDA is the gate keeper in this event based process, tasked with determining if the system meets the entrance criteria to proceed to the next phase of development. The following is a brief description of what occurs at each milestone [25, 54, 52, 51, 59]:

- *Milestone A:* The MDA approves a materiel development decision and grants formal entry into the development process. An Analysis of Alternatives (AoA) is conducted to determine which concept best satisfies the capability gap identified

المنارة للاستشارات
www.manaraa.com

in the JCIDS process. This phase is complete when a recommendation is made for technological development of the system concept identified at the end of the AoA.

- *Milestone B:* Ensures that a system is ready for production. This requires certification that all technology is mature enough for system-level development, the appropriate JCIDS documentation is approved, and all relevant funds are allocated.

- *Milestone C:* Represents the commitment to Low Rate Initial Production (LRIP), or procurement for those systems that do not require LRIP. If LRIP is deemed successful, a subsequent Full Rate Production (FRP) review will transition the system to full production status.

To support the program management decision-making process, credible and timely technical information that covers the entire system life cycle must be available to the decision-makers at each milestone [59]. This is particularly true of an OSA design concept, where technology will be repeatedly infused throughout the O&S phase to satisfy evolving requirements. Any forecast of evolution in requirements and technology are, however, inherently uncertain. This information must therefore contain an adequate representation of the development risk, as well as any contingency plans to mitigate this risk. As will be shown in the coming sections, this information is rarely available in the present acquisition development process. New approaches to gather and present this information to decision-makers will therefore be required.

2.2 Analysis of Alternatives

The challenge of designing solutions to meet demanding, and often uncertain, requirements at a reasonable cost is not unique to field of military acquisitions, nor is it a recent phenomenon. Conklin eloquently describes the fundamental nature of this

المنارة للاستشارات
www.manaraa.com

challenge [40].

> Any design problem is a problem of resolving tension between what is needed and what can be done. On the one hand, the process of design is driven by some desire or need - someone wants something new. The need might be expressed by a customer, or it may be a guess about what the market wants. The need or want is expressed in the language of what ought to be - what should be done, what should be built, what should be written.
>
> On the other hand, the process of design is constrained by resources - what can be done given the available resources such as time and money and the constraints imposed by the environment and the laws of science. Every need has a price tag - the process of design is about devising solutions that are feasible and cost effective.
>
> When an individual does design, she stands with one foot in each world. Moving back and forth between the two worlds, she tries to create a solution that joins the two polarities of design in an elegant way.

Acquisition professionals have refined a general process for joining these two polarities of design in the form of the Analysis of Alternatives methodology. An AoA is broadly defined as an analytical comparison of the operational effectiveness, cost, and risks of the proposed materiel solutions to gaps and shortfalls in operational capability. The outcome of this approach should therefore provide answers to the following overarching questions [135, 51].

- What alternatives provide validated capabilities?
- Are the alternatives operationally effective and suitable?
- Can the alternatives be supported?

المنارة للاستشارات
www.manaraa.com

- What are the risks (technical, operational, programmatic) for each alternative?
- What are the lifecycle costs for each alternative?
- How do the alternatives compare to one another?

2.2.1 Effectiveness Analysis

Effectiveness analysis is normally the most complex element of the AoA. Its goal is to determine the military worth of the alternatives through a bi-level decomposition. The first tier in the hierarchy is composed of the Mission Tasks (MT) describing the general actions to be performed, or the effects to be achieved by the system. These represent the first pole in Conklin's analogy of design, and are therefore formulated in the voice of the warfighter (e.g. obtain responsive intelligence surveillance and targeting information). The second tier of the hierarchy provides qualitative or quantitative measures of a system's performance that indicate how well it performs the corresponding task. These measures are referred to as Measures of Effectiveness (MoE), and are intended to represent raw quantities (e.g. percent of targets detected). Figure 8 depicts a notional MT/MoE decomposition for an Intelligence, Surveillance and Reconnaissance (ISR) capability requirement [135, 173].

It should be self-evident from the example in Figure 8 that alternative systems cannot be directly evaluated in terms of MoEs. Rather, the measurable properties of a system are, in the acquisitions vernacular, known as Measures of Performance (MoP). Representative examples of MoPs include range, velocity, mass, and weapon load-out [59]. These parameters serve as the inputs to Modeling and Simulation (M&S) analyses, which, along with documented constraints and assumptions, generate the MoE used to assess to overall military worth of the alternatives under consideration. This combination of MoP, M&S, constraints, and assumptions serves as the second pole of Conklin's design analogy. In addition, variations in assumptions and MoP values make it possible to investigate performance sensitivities in the robustness of

www.manaraa.com

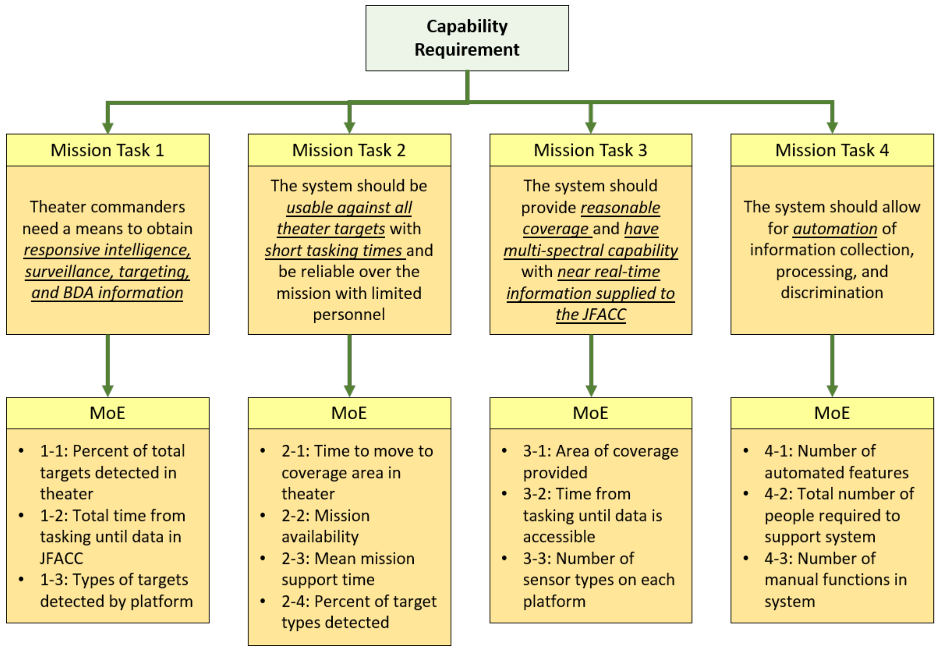


Figure 8: Notional Capability Requirement [135]

alternatives whose defining parameters are subject to uncertainty [135].

2.2.2 Cost Analysis

Cost analysis is usually conducted in parallel with effectiveness analysis, and is considered to be of equal importance. The core consideration of a cost analysis in the acquisitions context is the Total Life Cycle Cost (TLCC) of the system. Note that this is a significant departure from commercial ventures, which also consider sunk costs, break-even points, etc. Estimation of TLCC is usually combined with the results of the effectiveness analysis to perform a combined cost-effectiveness comparison. Elements of TLCC are generally grouped into one of the following categories [135, 51, 59]:

- Research, Development, Test, and Evaluation (RDT&E)
- Production (Low Rate Initial Production and Development)
- Operations and Sustainment (O&S)
- Disposal Cost

www.manaraa.com

There are several different methods for estimating and aggregating the system costs in each of the categories listed above. What follows is a brief synopsis of the most widely used approaches [135, 73, 56].

2.2.2.1 Analogy

Most programs evolve from previous efforts that have had new features added, or simply represent a new combination of components. Estimating costs by analogy uses this logic in order to infer the cost of a new program from observations of a similar program, with adjustments to account for differences in requirements. The adjustments are typically made through the use of scaling parameters to account for variations in size, performance, technology, or complexity. This method is advantageous since it can be used before detailed design requirements are known, can be implemented quickly and cheaply, and is simple to understand. However, this method is often criticized for relying on a relatively small pool of data points, extensive use of expert opinion, and detailed cost data for legacy systems that may not be readily available.

2.2.2.2 Engineering Build-Up

As the name implies, this method builds an overall system cost estimate by summing or "rolling up" detailed cost estimates performed at the component / sub-system level. Estimates typically consist of labor and material costs developed in consultation with the contractor's design team, which are then augmented with overhead and profit margins. This method is often applied during the detailed design or production phases, because the system design must be fairly stable and well defined for results to be reliable. As such, engineering build-up estimates are well regarded for being traceable, easily audited, and providing significant insight into the major cost drivers. On the other hand, this approach is typically expensive, time consuming, and not sufficiently flexible to allow for "what if" considerations. Further, it requires tremendous

للاستشارات www.manaraa.com

effort to remain up to date in the presence of production or process changes, and requires a separate analysis for each alternative under consideration.

2.2.2.3 Parametric

Parametric cost methods develop a statistical relationship between historical costs and their respective programmatic, physical, and performance characteristics. Such a relationship is known simply as a Cost Estimating Relationship (CER). The underlying assumption in developing a CER is that the same factors impacting costs in previous development efforts will continue to impact future costs in the same manner. It is therefore critically important that the factors used to develop these relationships (e.g. power, weight, lines of code, etc.) represent the principal factors driving the costs of previous programs. In addition, it is essential to have an adequate number of relevant data points in order to ensure that a CER passes the stringent tests of statistical significance necessary to provide confidence in the results.

These methods are very common in cost estimation because they can be developed early in the design process, are applicable at any level of design, and, unlike their engineering build-up counterpart, are easily adapted to consider "what if" questions or alternative designs. Further, these methods are able to assess the sensitivity of results or assumptions by perturbing the corresponding input parameters. The drawback of the CER, however, is the need to develop and maintain a database of detailed cost data, which must remain up to date. In addition, programs whose data points fall beyond the range of inputs used to generate the CER run the risk of significant error, as CERs are not well suited for extrapolation.

2.2.2.4 Expert Opinion and Extrapolation

The preceding methods are considered to be best practices for cost estimation, but there are instances in which none of these methods are applicable. Expert opinion can be useful in these scenarios, particularly when there is no historical data available to

المنارة للاستشارات
www.manaraa.com

generate a useful CER, and the concept lacks sufficient definition for an engineering build-up. Moreover, expert opinion can be advantageous when applied in conjunction with more rigorous methods to ensure that all variables and contingencies are considered. Caution should be exercised, however, as expert opinions are often considered to be too subjective and inaccurate to function as an authoritative estimate. This is particularly true when there is a potential conflict of interest between the experts available and the project under consideration.

The final category of cost estimation techniques are the extrapolation methods. Extrapolation methods use historical data to formulate statistical relationships, but, unlike CERs, the requirement for interpolation is disregarded in order to evaluate programs that lie outside the bounds of previous development efforts. Again, care should be taken when assuming the risk of extrapolation, as significant error is likely. In the event that such risk is unavoidable, however, it is best to pair extrapolation with expert opinion as a mitigation strategy.

2.2.3 Cost-Effectiveness Comparisons

Cost and effectiveness analyses provide an estimate of the TLCC and the MoP for all alternatives under consideration. In order to select an alternative, however, decision-makers must determine if the "value" provided by an alternative is worth its corresponding cost. This is not an uncommon problem. Challenges regarding resource allocation decisions for a portfolio of investment opportunities in new products or research and development efforts are ubiquitous in commercial ventures. However, commercial firms seeking to maximize profit in a competitive market have a significant advantage in that the broader market invariably communicates the the relative value of a good, service, or investment through the equilibrium relationship between supply and demand. Commercial firms are therefore better able to objectively compare the relative costs of developing an opportunity to the expected return on its investment

المنارة للاستشارات
www.manaraa.com

of resources when evaluating competing alternatives. On the other hand, there is no such mechanism to establish the fair market value of alternative defense systems the DoD wishes to acquire. The major cause of this distinction stems from the fact that the DoD is often the most important, and sometimes the only, customer to the defense contractors supplying such systems [59]. Consequently, government acquisition decisions directly impact the performance of its supplier base, and thereby distort prices [139]. Augustine summarizes the concept as follows [143]:

> On the surface, defense acquisition appears to have little in common with commercial acquisition. For starters, defense acquisition occurs in a monopsony[3]. Further, it is replete with mini-monopolies. (From how many places could one have purchased, say, an additional B-2?). Defense acquisition also operates in a governmental system that intentionally traded optimum efficiency for strong checks and balances - such as those implicit in separating Legislative and Administrative branches. Nonetheless, there are certain fundamentals of sound management which are applicable virtually everywhere, including in the defense acquisition process. They are just more difficult to apply in the government, where the stakes are higher, authority less hierarchical, and the spotlight much brighter.

Compounding this problem is the fact that national defense is a purely public good, which is defined by two properties. The first is *non-rival consumption*, where one customer's consumption of a marginal unit of the good or service does not preclude another's consumption of the same unit. The second characteristic is *non-excludability*, which requires that the good or service cannot be provided to an individual without simultaneously providing it to other. Economists have long studied this concept and, though there are arguments around peripheral concepts/definitions,

[3] A *monopsony* exists when there is imperfect competition, as only one buyer faces many sellers, or the opposite of a monopoly [146].

المنارة للاستشارات
www.manaraa.com

have come to the consensus that such goods cannot be valued in the same way as non-public goods [83]. This prompts non-profit organizations and agencies such as the DoD to adopt either a cost-effectiveness analysis or cost-benefit analysis to evaluate different alternatives. Cost-effectiveness analyses seek to identify and place a dollar value on the costs of a program, and then weighs those costs against the dollar value of program benefits. The net benefit provided by program is then determined by subtracting the total cost from the aggregate benefit [30]. The Achilles' heel of cost-benefit valuation lies in its explicit requirement to place a dollar value on the benefits provided by the program. It is not clear, however, how many units of "defense" are provided by a given program, and this is especially true when system operate as part of a greater system-of-systems to satisfy an operational or strategic military need. Moreover, if it were possible to determine a program's relative contribution to the nation's collective defense, one would still be left with the challenge of assigning a dollar value to concepts like freedom and security. Consequently, cost-benefit valuation is typically reserved for a limited number of situations, leaving cost-effectiveness analyses as the preferred methodology [59, 103].

Similar to cost-benefit, cost-effectiveness analyses seek to identify and place a dollar value on the costs of a program. The critical distinction is that the benefits of a program are left in the native units of "effectiveness", as specified by the appropriate MoE. Analysts then obtain a program's cost-effectiveness ratio by dividing the program cost by units of effectiveness. The challenge with this approach is that it may not be reasonable, or even possible, to combine multiple MoE into a single, aggregate measure. Further, there is some agreement within the acquisition community that relative, as opposed to the absolute valuation of cost-benefit methods, can be potentially misleading. In essence, this argument states that knowing the relative cost-effectiveness of a group of alternatives does not guarantee that the most cost-effective option justifies the investment of resources [59]. Levin summarizes this

المنارة للاستشارات
www.manaraa.com

argument [103]:

> That is, we can state whether a given alternative is *relatively* more cost-effective than other alternatives, but we cannot state whether its total benefits exceed its total costs. That can only be established through a cost-benefit analysis.

The conclusion from these observations is that cost-effectiveness is the best mechanism to evaluate acquisition programs, but it is difficult to determine when an incremental improvement in effectiveness is worth its corresponding cost. Acquisition decision-makers therefore use cost-effectiveness plots like the one shown in Figure 9 to manage these decisions. In addition to plotting data points of cost and effectiveness, designers are also concerned with the sensitivity of results to uncertainty in MoP, assumptions, etc. The sensitivity of results, assuming a pre-specified confidence interval, is indicated by the box surrounding each point.

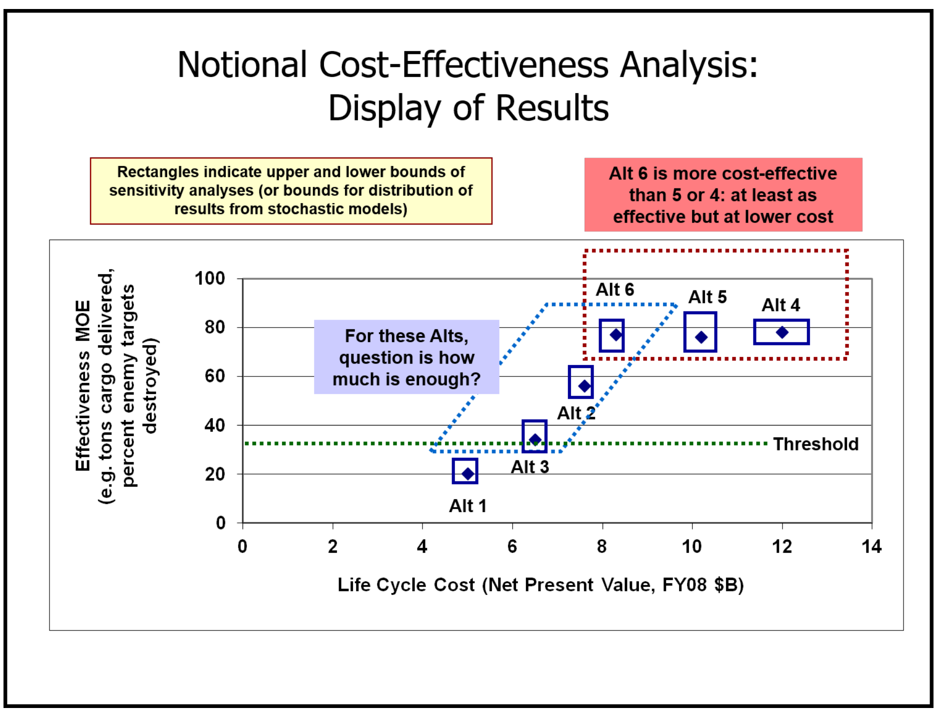


Figure 9: Notional Scatter Plot of Cost vs. Effectiveness [100]

المنارة للاستشارات
www.manaraa.com

From Figure 9, decision-makers can safely conclude that Alternative One is a poor choice because it falls below the minimum required threshold for effectiveness. Alternative Three requires closer scrutiny. It is the lowest cost of all Alternatives that are expected to meet the threshold requirement, but there is a considerable risk that the real system will fall below this threshold. Additionally, Alternatives Four and Five can be excluded from further consideration because the remaining alternatives provide at least as much effectiveness, but they do so at a lower cost. No clear argument distinguishes Alternative Two and Alternative Six, however, as Alternative Six would be chosen if the increase in effectiveness is deemed to be worth the additional cost. Decision-makers must make this distinction.

2.3 Methods for Open System Architectures

The previous section established that determining the "value" of a traditional, closed system architecture can become problematic. OSA pose a more complicated decision space for acquisition decision makers because the system level MoP are intended to change over time in response to evolving requirements. Changes in MoP would then propagate through M&S - along with time variations in assumptions, constraints, and threshold effectiveness - to cause changes in MoE. This is equivalent to adding a time dimension to the cost-effectiveness scatter plot. However, the TLCC and the effectiveness at a given point in time depend on the upgrade decisions made between the present and the point in time being considered. It is not clear at this time how such alternative upgrade plans could be generated using current practices. Moreover, it even less clear how the considerable uncertainty inherent in forecasting the evolution of technology and system requirements could be incorporated into this visualization. What is clear is that if traditional methods struggle with determining when additional cost is worth additional effectiveness, then they are not well structured to determine if added effectiveness provided over time is worth the additional resources in the

المنارة للاستشارات
www.manaraa.com

present.

The shortcomings in applying traditional decision structures to OSA is not an original observation. Significant research has been conducted within the acquisition community over the last two decades aimed at resolving this discrepancy. These methods can be grouped into one of two categories, depending on whether they are applied at the architectural level or the module level.

2.3.1 Architectural Level Methods

Early research conducted by the acquisitions community on OSA design led to the realization that the new paradigm requires careful consideration of numerous factors the were not addressed by existing SE methods. Moreover, they observed that failing to properly plan for and execute these factors could undermine the utility of the open design, leading to a system that was more costly and less effective than one developed under traditional methods. The challenge with accommodating these factors stems, in large part, from the fact that they range across the business, technical, and management dimensions of any potential development program. For example, a design cannot be considered "open" if contractual language for securing IP rights *(business)* and configuration management practices for key interfaces *(technical)* are not established early in the development process. Further, these systems mean little if training and oversight are not established to ensure accountability *(management)*. The reach and complexity of these factors therefore made it difficult to specify a single, holistic set of standards and "best practices" for open development programs [51].

The DoD empanelled the Open System Joint Task Force (OSJTF) in 1994 to confront this challenge. Nearly 10 years later, the OSJTF concluded that the breadth of considerations facing PMs precluded the use of a uniform approach to OSA design. Rather, they proposed a set of guiding principles shown in Figure 10. In addition, the

المنارة للاستشارات
www.manaraa.com

observed principles were decomposed into a set of 60 programmatic and technical indicators to help developers assess the extent to which OSA principles are implemented in an acquisition program. A complete list of indicators are provided in Appendices A and B [145]. Examples of these indicators include the following:

- **Programmatic:** To what extent have program requirements been analyzed, and refined as needed, to ensure that design-specific solutions are not imposed?
- **Programmatic:** To what extent does the program plan include lifecycle support and funding for open architecture elements?
- **Technical:** To what extent does the system's architecture exhibit modular design characteristics?
- **Technical:** To what extent is the system's architecture capable of adapting to evolving requirements and leveraging new technologies?

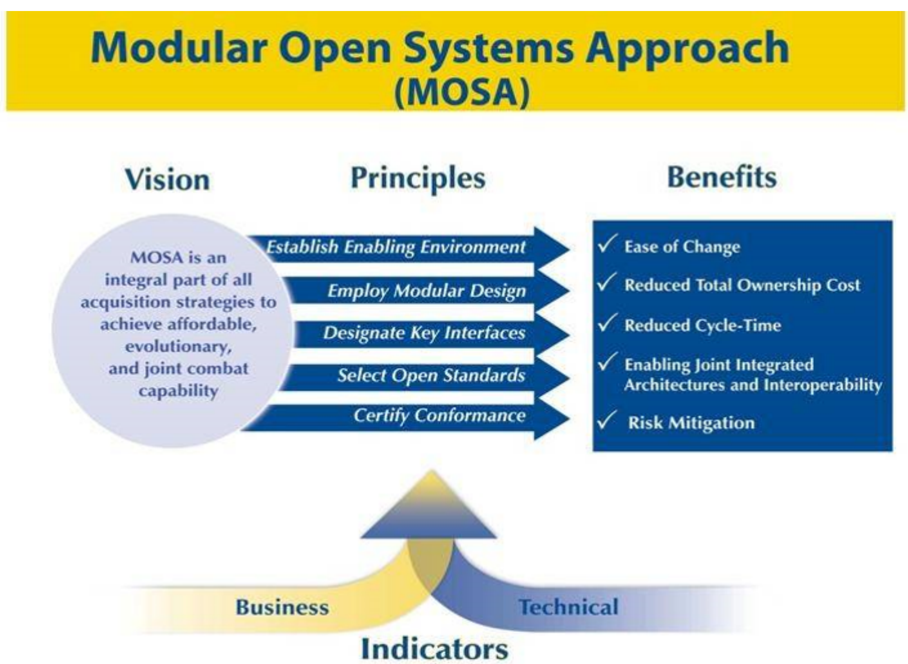


Figure 10: OSJTF Principles for Modular Open System Architectures [145]

المنارة للاستشارات
www.manaraa.com

The responses to all implementation questions are qualitative in nature, and broadly consist of the following alternatives: *None, Little Extent, Moderate Extent*, and *Large Extent*. These responses are tabulated in an Excel-based tool known as the Open Architecture Assessment Tool (OAAT). The results are then evaluated with the Open Architecture Assessment Model (OAAM) embedded in the worksheet to produce a score of 0-100% for the level of "openness" the program possesses with respect to its programmatic and technical practices. These normalized scores are then plotted on the Open Architecture Maturity Matrix (OAMM) provided in Figure 11 to ascertain a description of the program's current state relative to other programs. Table 1 provides the linkage between the OAMM and this description [51, 145].

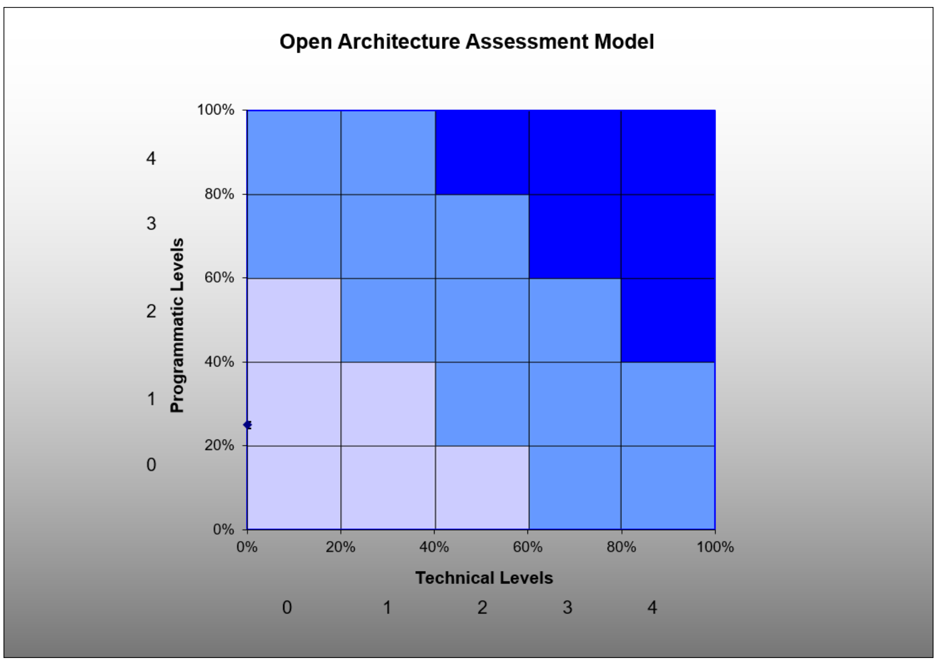


Figure 11: Open Architecture Maturity Matrix [145]

2.3.2 Module Level Methods

The OAAT/OAAM approach discussed in the previous section is intended to ensure the the necessary conditions are in place for an OSA to be successfully developed.

www.manaraa.com

Table 1: Qualitative Interpretation of OAMM Results [145]

Level	Programmatic Description	Technical Description
0	Isolated	Closed
1	Connected	Layered
2	Migrating to Openness	Layered and Open
3	Common	Common
4	Open and Net-Centric	Enterprise

They do not, however, provide guidance as to how the designer should identify which system elements should be open to further upgrades. The two major methods in place to manage such decisions are the Risk Assessment approach, and the Key Open Sub-System (KOSS) methodology [51, 166].

The purpose of classical risk assessment methods is to evaluate the significance of future uncertainties in achieving program performance goals and objectives within defined cost, schedule and performance constraints [53]. Figure 12 depicts how traditional risk assessment methods and techniques can be applied to an open system context [51]. Implementation begins with a system-level decomposition of physical components, where any exchange of information between components is depicted as an edge connecting two nodes at the same level of the hierarchy. Components at the lowest level of abstraction are then evaluated independently by considering the likelihood they will require modification or replacement due to rapid evolution of related technology, changes in requirements, or low mean time to failure. The likelihood and severity of these events are then evaluated against the classical risk reporting matrix (lower portion of Figure 12) to qualitatively determine the overall risk of integrating the component into the greater architecture. Low risk components are integrated and high risk components are opened, though it is not immediately clear how intermediate components should be evaluated. Once the key interfaces are identified, the final step step is to recursively identify interfaces in higher level modules that must be opened in order to ensure that the component level interfaces are accessible [51].

المنارة للاستشارات
www.manaraa.com

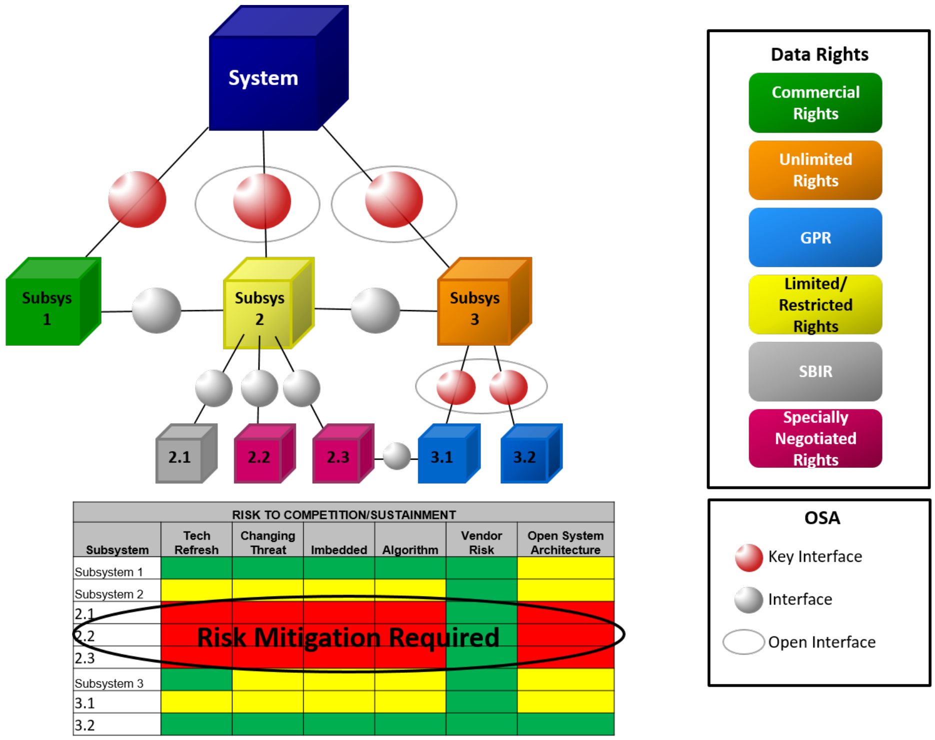


Figure 12: Risk Assessment Process for OSA [51]

Though the Risk Assessment method is simple and cost efficient, it fails to provide explicit guidance as to how the likelihood and severity of replacement/modification should be ascertained. The second method under consideration, KOSS, is another Excel-based tool developed by the U.S. Navy Air Systems Command (NAVAIR) and various industry partners to address this shortcoming. Specifically, KOSS is intended to aid PMs by providing greater transparency in determining why a given component is volatile, and what return on investment would be expected if the component were opened to future upgrades. The basic algorithm is provided in Figure 13 [166].

As with the risk assessment approach, the first step in this process is to physically decompose the system into a set of basis components. In addition, the expected time variation in requirements is converted into a Capability Road-Map (CRM), which specifies how the threshold MoEs are expected to change over discrete steps in time. A given component is evaluated against this CRM to determine if a change is needed

www.manaraa.com

in order to meet the various thresholds. Results are then summed using a binary operator, 1 if a change is required and 0 if no change is required, and recorded as a CRM score. Next, the component's likelihood of obsolescence, cost to replace/modify, and relative weapon system capability improvement are assigned a qualitative score of Low, Medium, or High. Qualitative scores are assigned and mapped to numeric values in accordance with the guidance in Table 2. The first step in determining the Return On Investment (ROI) of opening a given component is to sum the CRM score with the corresponding obsolescence score. The result is defined as the relative rate of change, which is then multiplied by the relative cost of change to determine the rate at which resources must be expended to keep the component in line with the CRM. Finally, the rate of resource allocation, identified as "OA Applicability" in Figure 13, is multiplied by the relative capability improvement to determine the overall ROI to the warfighter [166].

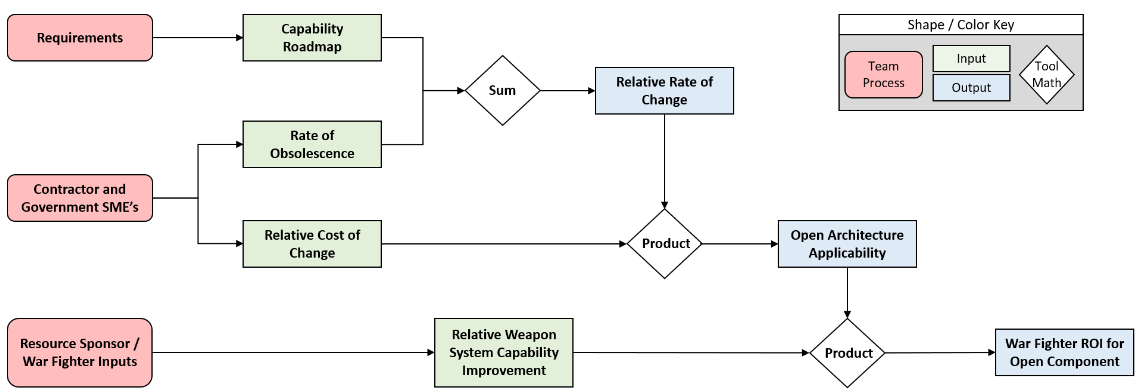


Figure 13: Key Open Sub-System Methodology [166]

The output of the KOSS model provides two useful results to the user. First, analysts and decision-makers can use the ROI scores to determine where constrained resources should be allocated to achieve the greatest impact. Second, the reduced capability roadmap, one which includes only those components intended to be upgradeable, provides planning guidance as to when components will be upgraded and how much that upgrade is expected to cost. The evolutionary path defined by this

المنارة للاستشارات
www.manaraa.com

Table 2: Qualitative Interpretation of KOSS Input Levels [166]

Rating	Obsolescence	Score	Change	Score	Weapon	Score
Low	Low probability to go obsolete within time period	0	Cost of change less than 10%	1	Infrastructure changes	1
Med	Probability of one change within time period	1	Cost of change between 10% and 33%	3	Moderate improvement. Evolutionary change	2
High	Probability of two or more changes within time period	2	Cost of change greater than 33%	6	Large improvement. Revolutionary change	4

information is referred to as the system's *Technology Refresh Plan* (TRP). TRP's are considered to be a critical component to the lifecycle management of upgradeable systems [42].

2.3.3 Gap Analysis

In summary, each of the the methods previously discussed facilitate the design of OSA systems in different ways - OAAT/OAAM sets the conditions for successful development, Risk Assessment / KOSS approaches search for optimal configurations, and the AoA methodology provides a comparative framework. Yet, the fact remains that implementation of OSA design principles remain limited due, in large part, to a lack of confidence that existing methods can reliably determine whether the potential benefits of an OSA justify the corresponding cost and risk. This criticism is not unwarranted. The Risk Assessment approach yields qualitative insight on the risk drivers underlying a given decision, but it does not provide a consistent decision mechanism for alternatives exhibiting "Moderate" risk. In addition, it is not clear how the TRP guidance (e.g. when should technology be infused, how much will it

المنارة للاستشارات
www.manaraa.com

cost, etc.) necessary for lifecycle management can be inferred from these qualitative results. As such, this approach is bested suited as a screening process to reduce the set of all components to a sub-set of those that deserve closer scrutiny.

The KOSS model is well regarded in acquisition circles for addressing the shortcomings of the Risk Assessment approach, though the authors acknowledge that the model is a starting point for further research [166]. The challenge to practical application of the KOSS model is the assumption that Subject Matter Experts (SME) can determine, with absolute certainty, which technology will be chosen at future points in time, the corresponding changes to component MoPs and system MoEs, and the cost associated with the infusion of this technology. Analysts and decision-makers recognize, correctly, that these factors are subject to considerable uncertainty, but the treatment of this uncertainty in the KOSS model is opaque. Moreover, much of the potential of OSA designs lies in the fact that the decision-maker is not committed the single, rigid development path framework espoused by the KOSS model. Rather, decision-makers have the flexibility to alter decisions overtime in order to capitalize on technological advances that are difficult, if not impossible, to predict.

Overcoming the barriers to successful implementation identified in Chapter One will therefore require a more rigorous approach. One that is quantitative rather than qualitative, based on empirical analysis rather the SME opinions, and capable of articulating risk in terms of the balance between uncertainty in forecasts and flexibility in decision-making.

2.3.4 Research Objective

The objective of this research is to develop a methodology to evaluate Open System Architectures in terms of their expected costs, benefits, and risks in such a manner that they are directly comparable to traditional closed system architectures. To determine how well this methodology aids decision-makers in the identification of a

المنارة للاستشارات
www.manaraa.com

cost-effective, evolutionary strategy for open system development, the methodology should be judged against criteria representing the stated needs of the acquisition community. Therefore, the following should be addressed:

- Must take into account the likely evolution of commercial technology and system requirements [14, 75, 70]
- Must take into account the uncertainties inherent to forecasts of future events [55, 60]
- Must take into account the potential of flexibility to alter decisions as a hedge against uncertainty [14, 55, 60]
- Must result in a codified Technology Refresh Plan to aid in lifecycle planning [42]
- Must minimize the trepidation of acquisition professionals in pursuing the OSA paradigm by maximizing the use of existing acquisition methods and techniques [135, 60]
- Must ensure that the advantages and disadvantages of alternative development strategies are presented in a clear and unbiased manner, and that it depicts the analysis results, understandable interpretations, and defensible recommendations [135, 51]

Currently, estimates of the value added by imposing open architecture constraints are limited in use. The numerous, uncertain factors contributing to this valuation are either ignored, or treated as qualitative parameters to be set and aggregated by SMEs. *The key difference that sets this methodology apart from other research is its consistent, deliberate, and traceable approach to evaluate the system and its potential for evolution in parallel. In so doing, it will become possible to directly compare open and closed architectures.* The hope is that providing this capability will

المنارة للاستشارات
www.manaraa.com

allow for more informed decisions as to when and where open architecture constraints are appropriate, which will, in turn, lead to greater implementation of the OSA paradigm.

www.manaraa.com

CHAPTER III

RESEARCH QUESTIONS

3.1 Research Question One: Systems Engineering Requirements

A summary of the observations leading to the first research question are:

O1a. The underlying engineering principle for Open Architecture Systems is that physical modularity allows open components to be upgraded at a low cost. This ability to efficiently upgrade components is desirable because it allows evolving, uncertain requirements to be satisfied later in the system's life cycle, where uncertainty is reduced and technology is less expensive and/or possesses better performance.

O1b. Use of custom components is believed to be expensive, as the government is obligated to pay all fixed costs associated with RDT&E, production set-up, and overhead. It is therefore desirable to use of commercial components wherever possible, since the fixed costs of commercial products are amortized across a much larger base.

O1c. Imposing excessive requirements early in the system's design process (e.g. requiring high confidence of satisfying terminal requirements at Initial Operating Capability) biases the analysis process to prefer custom components.

O1d. System designs with insufficient and/or sub-optimal modular partitioning schemes are likely to overlook potential opportunities to infuse commercial technology. This represents a second mechanism of analytical bias towards custom, integrated architectures.

المنارة للاستشارات
www.manaraa.com

Based on observations O1a. through O1d., the focus of the first research question becomes:

RQ1. What are the necessary requirements to ensure that early Systems Engineering analyses do not bias results toward integrated architectures with custom components?

3.2 Research Question Two: Technology Refresh Planning

A summary of the observations leading to the second research question are:

O2a. The government has little to no ability to direct the process of technological evolution in commercial markets. At the same time, this process is the mechanism through which open systems efficiently satisfy future requirements.

O2b. Decision-makers will not accept that an open architecture will satisfy future requirements at a lower cost on faith. A technology refresh plan must be presented along with an open design to define when the various components will be upgraded, and how much these upgrades will cost.

O2c. There is a fundamental mismatch between observations O2a and O2b; the government has no control over the process of technological evolution, but must plan as though this uncertain progression were known.

Based on observations O2a. through O2c., the focus of the second research question becomes:

RQ2. What is an appropriate method to develop an optimum technology refresh plan that leverages technological evolution to efficiently meet evolving requirements?

المنارة للاستشارات
www.manaraa.com

3.3 Research Question Three: Balancing Uncertainty and Flexibility

A summary of the observations leading to the third research question are:

O3a. Evolution in technology and requirements are predictions as to how the future will unfold. Such predictions are inherently uncertain.

O3b. Existing methods provide an opaque, qualitative treatment of uncertainty based on the opinions of SMEs. Acquisition decision-makers are reluctant to accept the results of such methods.

O3c. Uncertainty in predictions is mitigated by the fact that decision-makers are not locked into a rigid technology refresh plan. Previous decisions can always be adapted in response to new information.

O3d. The flexibility to alter decisions serves as a hedge against the uncertainty inherent to forecasting methods.

Based on observations O3a. through O3d., the focus of the third research question becomes:

RQ3. What theories or methods can be used to convey the impact of uncertainty in requirements and technological evolution in the presence of the decision-maker's flexibility to alter technology refresh plans as new information is provided over time?

3.4 Research Strategy

The stated objective of this work is to develop a methodology that resolves observable gaps in existing OSA design methods. While the aforementioned research questions must necessarily be addressed to achieve this end, they are not sufficient. What remains to be determined is how the answers to these questions inform a process model

www.manaraa.com

that acquisition professionals can use to manage the complexities of open architecture design. Bridging this divide will require some form of predefined structure for the proposed methodology.

A direct approach to provide this structure this would attempt to map the research questions to specific activities in one of the existing methods (i.e. KOSS or Risk Assessment), at which point the activities could be modified to correct for the perceived deficiencies. This strategy would provide a simple way to focus the coming analysis, but it was ultimately determined to be impractical on two fronts. First, existing methods are based on deterministic, qualitative data provided by SME's. These are the precise requirements that must be removed to meet the stated needs of the acquisition community, which implies that the underlying premise of these methods, rather than their specific activities, must be modified to satisfy the research objective. Second, no existing method possesses steps that specifically address all of the considerations identified in the research questions. Thus, even if the underlying premise of existing methods could be adapted, additional activities must still be infused into the process model.

Considering the magnitude of modifications required to adapt one of the existing methods, it seems more appropriate to form a "bottom-up" strategy around a more general decision-making template. An example of such a template is found in the Georgia Institute of Technology's Integrated Product and Process Development (IPPD) framework depicted in Figure 14 [149]. This process was explicitly developed to define, measure, and evaluate technology under the Design for Affordability paradigm, and is therefore well suited to the objective of this work. Moreover, Sharma demonstrates that other well regarded decision support frameworks, to include NASA's trade study process, are subsumed under the more generalized IPPD model [152]. This work will therefore apply the IPPD template as the baseline decision-making framework for the proposed methodology.

www.manaraa.com

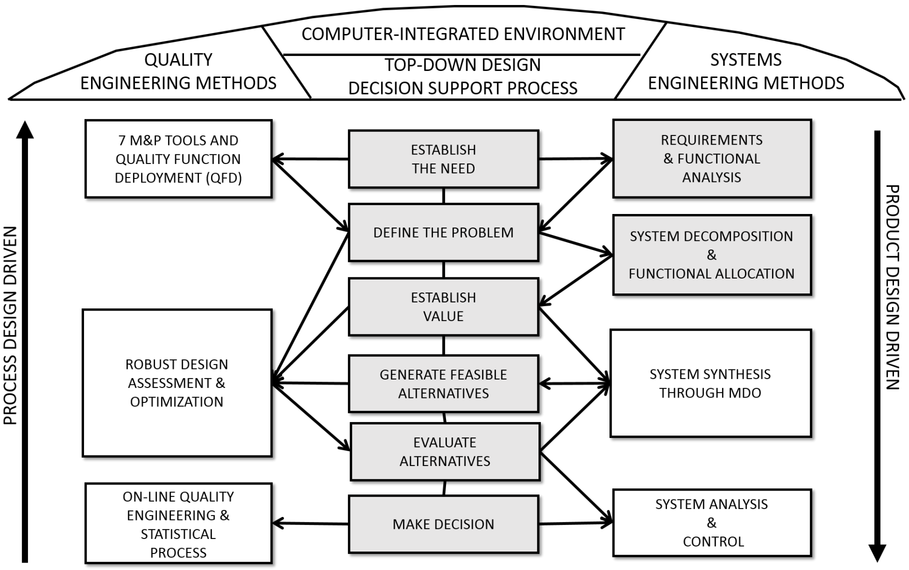


Figure 14: Integrated Product Process Development (IPPD) Framework [149]

Figure 14 demonstrates that the IPPD model is a composition of three complimentary fields, where individual fields are organized into columns. The primary contribution of the IPPD framework is the Top-Down Decision Support Process occupying the center column, which provides the unifying structure to integrate the most significant methods from both Systems and Quality Engineering. This decision support process will serve as the principal framework used to structure the proposed methodology. In addition, many of the challenges identified Research Question One pertain to the SE techniques occupying the right column of Figure 14. As such, these steps will either be included as explicit elements, or integrated with other elements of the top-down process model.

With this in mind, the steps for the general decision-making process under consideration in this research are defined as follows:

1. Establish the Need: Specify the purpose of pursuing an open design concept in terms of objectives and requirements

2. Modular Decomposition: Identify which portions of the system architecture contribute to the stated objectives / requirements

المنارة للاستشارات
www.manaraa.com

3. Define the Problem: Determine how the underlying technology of components identified in Step Two are expected to progress, and how that progression satisfies to the needs established in Step One
4. Establish Value: Establish measures of performance to compare alternative refresh strategies for the relevant component set
5. Generate Feasible Alternatives: Develop an automated procedure to identify feasible refresh plans warranting further consideration
6. Evaluate Alternatives: Assess the performance variability of alternatives in the presence of both uncertainty and flexibility
7. Make Decision: Present alternative plans to decision-makers in such a way that cost-effectiveness comparisons can be made

An abridged version of these descriptions is provided in Figure 15, along with their respective relationships to the stated research questions. Chapters Four, Five, and Six will address each of the research questions in the order in which they were presented. In addition, each chapter will begin with an initial analysis that decomposes its respective research question into more detailed sub-questions that address specific modules. A targeted literature review will then consider methods available to satisfy the requirements associated with the corresponding module, and Figure 15 will be updated as decisions are made. The methodology will be deemed to be "complete" once all modules in Figure 15 have been paired with their corresponding method. The final step will then verify that the proposed methodology satisfies the original research questions.

المنارة للاستشارات
www.manaraa.com

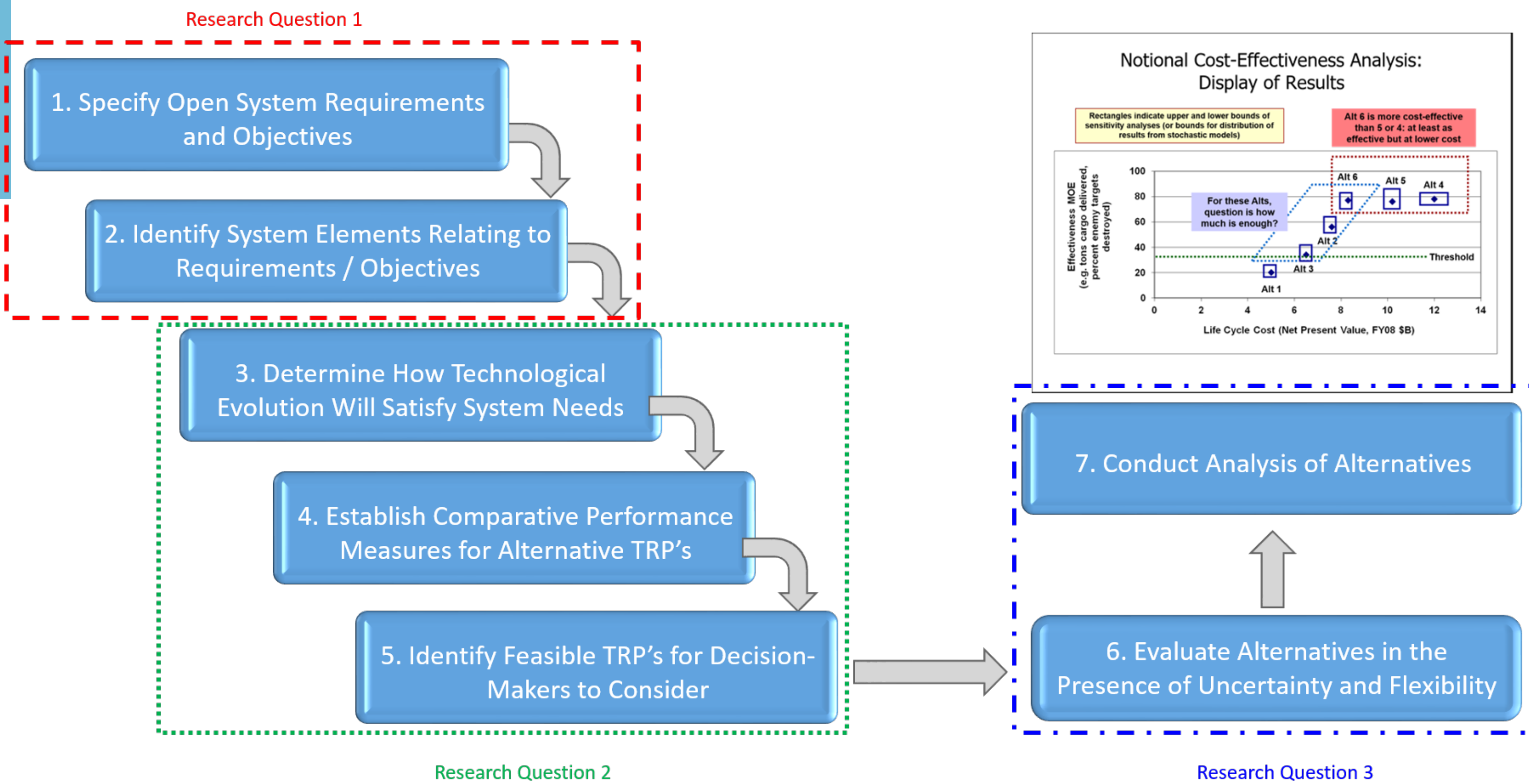


Figure 15: Research Strategy: Relation Between Research Questions and Objectives

www.manaraa.com

CHAPTER IV

REQUIREMENTS AND FUNCTIONAL ANALYSIS

The first research question follows from observations drawn from the OAAT/OAAM approaches. These methods are intended to ensure that the conditions for successful development of open systems are in place at the beginning of a development cycle. Failure to account for these considerations will undoubtedly bias the outcome of later analyses towards the selection of an integrated architecture with custom components. With that in mind, Research Question One can be refined by noting that Observations O1c. and O1d. represent two distinct bias mechanisms - improper requirements and improper modular partitioning. Looking back at the existing methods, it is clear that the KOSS model dealt with the former mechanism by mandating that requirements were defined in terms of a CRM documenting the time variation in threshold effectiveness. However, the KOSS model tends to be criticized for failing to provide a rigorous treatment of uncertainty. It is therefore appropriate to exercise the same CRM assumption for this work, with the caveat that some mechanism must be used to elicit uncertainty from the war fighter. On the other hand, it is not immediately clear from the research up to this point how one could ensure that the optimal modular partitioning scheme is identified early in the design process. These additional observations result in the following assumption, as well as a refinement of Research Question One.

المنارة للاستشارات
www.manaraa.com

Assumption

The war fighter specifies system requirements in the form of a Capability Road Map, which dictates the change in requirements over time along with a quantitative assessment of the corresponding uncertainty.

RQ1.1 What theories and methods are appropriate to elicit quantitative estimates of requirement uncertainty from the war fighter in order to properly develop a Capability Road Map?

RQ1.2 What theories and methods can be applied to ensure a design possesses sufficient modularity to maximize the added value of open architecture constraints?

The remainder of this chapter will focus on identifying academic methods in the literature with the potential to satisfy the needs of these requirements.

4.1 Eliciting Uncertainty in Requirements

The raison d'être for an OSA is its ability to efficiently satisfy evolving, uncertain requirements. This uncertainty stems, in large part, from the ambiguity intentionally incorporated into the DoD's strategic development road maps. For example, the Navy's Information Dominance Road Map over the 2020-2028 time-frame specifies the following objective [133]:

> Meet the growing demand coming from new Signal Intelligence (SIGINT) and ocean-based sensors, as well as higher resolution persistent sensors (including Full Motion Video (FMV) coming from space-based systems and multi-spectral sensors.

The logical question that follows from this objective statement is *how much additional demand will be required, and when will these requirements be imposed?* This

المنارة للاستشارات
www.manaraa.com

proves to be a difficult question to address in a consistent manner, as the requirements and their time-line depend not on the system itself, but on the parallel projects to develop more capable SIGINT sensors, FMV devices, and data analytics algorithms. It is unreasonable to assume that the PM and their staff will have the capability to review all relevant, external programs in order to create a realistic estimate of requirement growth. To resolve this challenge, this work will posit the existence of a team of Functional SME's whose sole purpose is to consolidate and synthesize this information into a CRM with the following format:

- When are requirements expected to increase?
- How much are the requirements expected to increase by?
- How confident are the Functional SME's in these estimates?

Figure 16 provides a depiction of what these questions mean for the design team developing the OSA platform that will be expected to host these improved payloads. The purpose of this section is to formulate a method to consistently identify and gather the information necessary to formulate this dimension of the problem.

4.1.1 Identifying and Modeling of Uncertainty

The fundamental concepts of truth, knowledge, and certainty have been an area of intense philosophical study since the early days of ancient Greece. The same can be said for the complimentary concepts of fallacies, confusion, ambiguity, uncertainty, etc. Several techniques have been developed to quantify the various ways in which knowledge can be imperfect during the design process. These methods include the popular approaches of Fuzzy Logic, Evidence Theory, Possibility Theory, and Monte Carlo probabilistic modeling. Selection of a given technique is not an arbitrary decision, as there are often subtleties as to what type of imperfection is truly being captured by each method. Recognizing this confusion, Ayyub created the taxonomy

www.manaraa.com

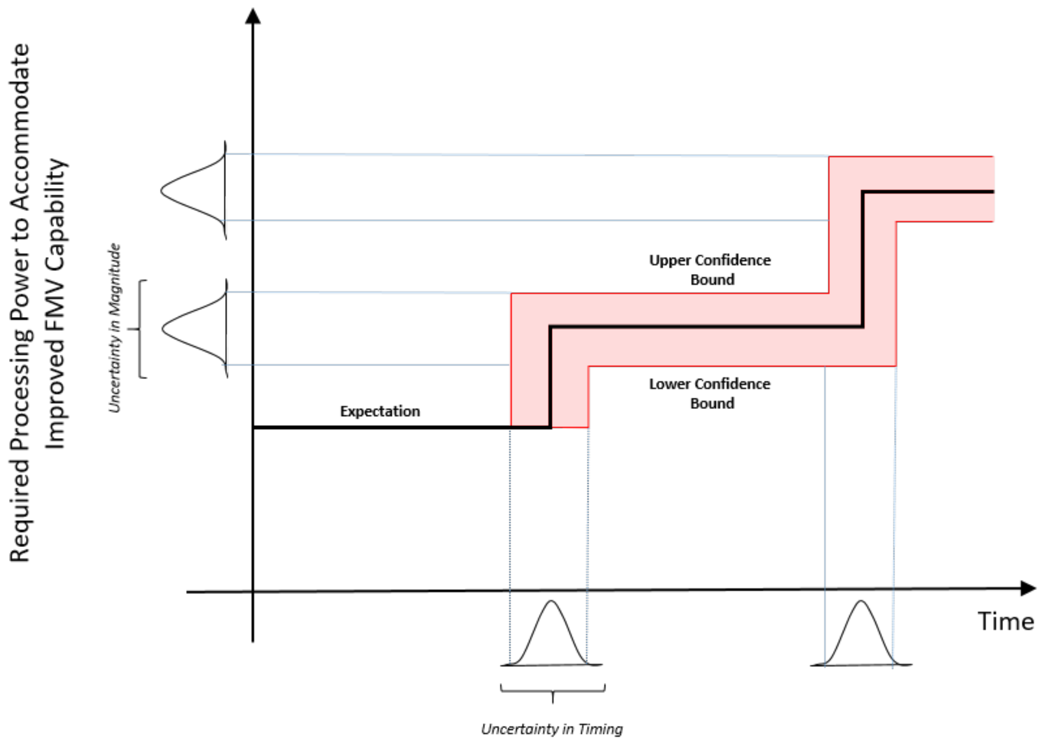


Figure 16: Uncertain Requirement to Accommodate Strategic Road Map

provided in Figure 17 to provide a more precise treatment of what drives "ignorance" in design. He also went on to conduct a review of the aforementioned methods for modeling this ignorance in the context of this taxonomy, which resulted in the mapping provided in Table 3 [11].

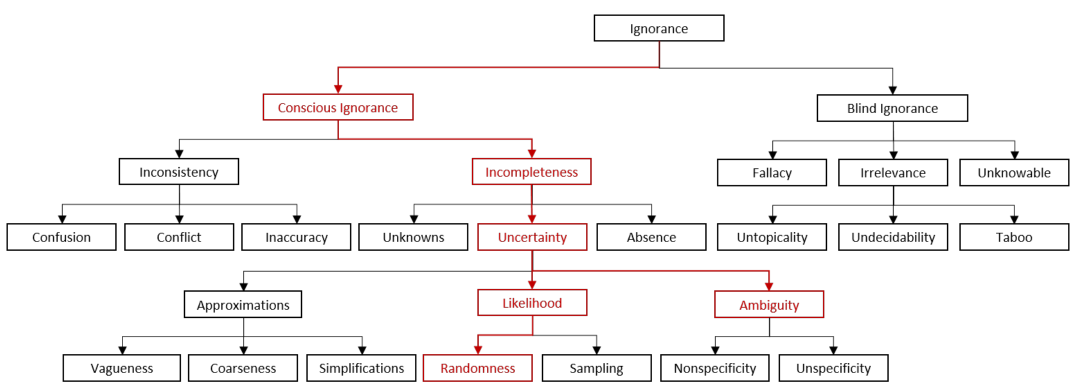


Figure 17: Ignorance Hierarchy [11]

Consider again the example taken from the Navy's Information Dominance Road

www.manaraa.com

Table 3: Theories to Model and Analyze Incomplete Information [11]

Theory	Confusion & Conflict	Inaccuracy	Ambiguity	Likelihood	Vagueness	Coarseness	Simplification
Classical Sets							
Prob. & Stat.							
Fuzzy Sets							
Rough Sets							
Evidence Theory							
Possibility Theory							
Monotone Measure							
Interval Prob.							
Interval Analysis							

المنارة للاستشارات
www.manaraa.com

Map. It would appear that there are two drivers of "ignorance" in the formulation of time-varying requirements for an OSA platform. First, the outcome of all parallel development projects is inherently unpredictable. In the language of Ayyub's taxonomy, this unpredictability of outcomes is known as "randomness". Second, there is no guarantee that planned development programs will actually come to fruition. Unplanned funding shortages or a failure of the underlying technology to reach maturity may lead to the outright cancellation of a program. Thus, even if it were possible to determine, with complete certainty, the requirements that *would be* imposed by parallel development efforts, the combinatorial uncertainty surrounding which programs succeed or fail would remain. According to Ayyub, this possibility for multiple outcomes is ambiguity. These branches of the ignorance heirarchy are highlighted in Figure 17. With this in mind, only one entry in Table 3 was noted to directly address Randomness and Ambiguity - Probabilistic Methods [11].

Probabilistic methods are unified in their interpretation of uncertainty in terms of Random Variables (RV). Random variables are defined by their respective *Probability Density Function* (PDF), $f_X(x)$, with the property that the probability of X occurring between two values, x_1 and x_2, is given by the integral of the PDF. This expressed by Equation 1.

$$P(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} f_X(X)dx \tag{1}$$

Alternatively, a RV can also be defined by the *Cumulative Distribution Function* (CDF), which returns the probability that X is less than or equal to a given value. These two forms, PDF and CDF, are equivalent expressions in the sense that the PDF can be defined as the derivative of the CDF, and the CDF can be defined as the integral of the PDF over its corresponding support. This is expressed by Equation 2

المنارة للاستشارات
www.manaraa.com

$$F(X) = P(X \leq x) = \int_{-\infty}^{x} f_X(X)dx \tag{2}$$

4.1.2 Parametric Definitions of Uncertainty

Mathematicians have derived several useful probability distributions over the years to capture commonly observed phenomena. The most direct method for a Functional SME to express their views on the uncertainty embedded in their expectations of requirement growth is to simply choose one of these distributions. Table 4 provides a list of several common distributions. If this is possible, then the only remaining challenge is to define the parameters and/or support for the corresponding distribution. Table 4 provides a list of several common distributions that a SME could potentially draw from.

Table 4: Common Probability Distributions [165]

Name	*Functional Form*	*Support*	*Parameters*
Uniform	$f(x) = \frac{1}{b-a}$	$x \in [a, b]$	$a, b > 0, a \neq b$
Exponential	$f(x) = \frac{1}{\theta}e^{-\frac{x}{\theta}}$	$x \geq 0$	$\theta > 0$
Gamma	$f(x) = \frac{1}{\Gamma(\alpha)\theta^{\alpha}}x^{\alpha-1}e^{-\frac{x}{\theta}}$	$x \geq 0$	$\alpha \geq 1, \theta \geq 0$
Normal	$f(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{\frac{(x-\mu)^2}{2\sigma^2}}$	$x \in \Re$	$\mu \in \Re, \sigma > 0$
Lognormal	$f(x) = \frac{1}{\sigma x\sqrt{2\pi}}e^{\frac{(ln[x]-\mu)^2}{2\sigma^2}}$	$x \in (0, \infty)$	$\mu \in \Re, \sigma > 0$
Triangular	$f(x) = \begin{matrix} \frac{2(x-a)}{(b-a)(b-c)} \\ \frac{2(b-x)}{(b-a)(b-c)} \end{matrix}$	$\begin{matrix} a \leq x \leq c \\ c \leq x \leq a \end{matrix}$	$a < c < b$
Weibull	$f(x) = \frac{\beta}{\eta}\left(\frac{x-\gamma}{\eta}\right)^{(\beta-1)} e^{-\frac{x-\gamma}{\eta}^{\beta}}$	$x \geq 0$	$\beta, \eta > 0, \lvert\gamma\rvert < \infty$

Ayyub notes, however, that simply choosing a distribution and its corresponding parameters may not be feasible for most real world examples. This statement follows from the observation that the parameters of common distributions either have little physical significance (e.g. Weibull shape parameters), or represent a physical process

المنارة للاستشارات
www.manaraa.com

with little direct applicability to the complex problems that require expert opinion. He argues that greater success has been found in estimating distributions in the context of their CDF, as opposed to their PDF. This is achieved through the method of *probability bounds* [11].

Probability bounds provide a consist method to address problems in which there is uncertainty in the probability of outcomes. The basic concept is to elicit CDF-based observations from a given expert. For example, the expert may be confident in expressing the minimum (e.g. 50) and maximum (e.g. 100) values that the RV can take, but is unwilling or unable to comment on the distribution of outcomes between those data points. Figure 18 depicts the probability bounds on the CDF that would arise from these data points, as well as a set of possible CDF's corresponding to distributions in Table 4 that would satisfy these constraints.

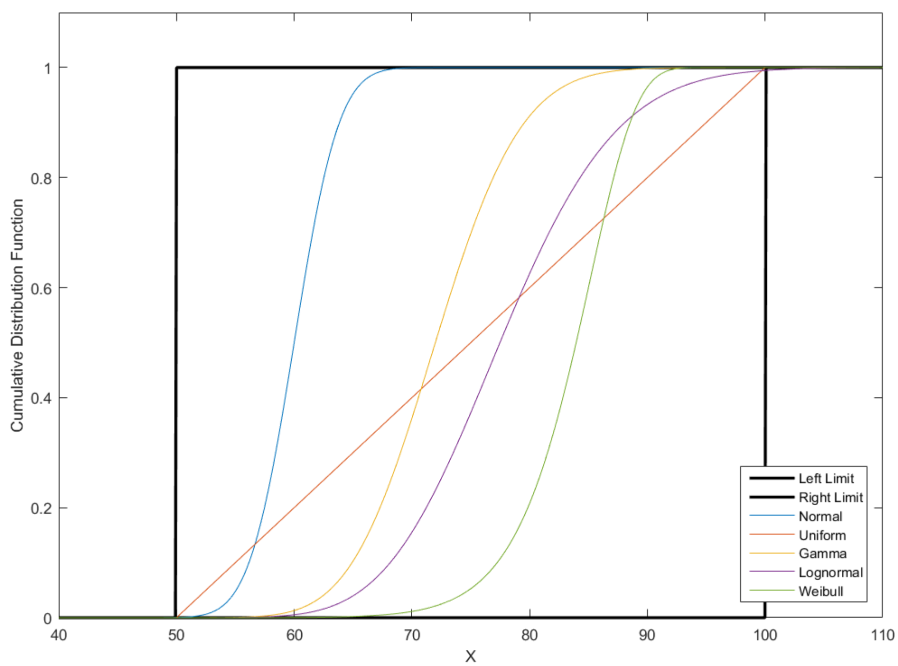


Figure 18: CDF Bounds for Minimum and Maximum Constraints

It is clear from Figure 18 is that there are a great many distributions that can

www.manaraa.com

satisfy these basic constraints, which may make it difficult to settle on a specific functional form and/or parameter set. However, the set of possible RV's can be dramatically reduced by adding additional constraints to the boundary set. Figure 19 demonstrates this point by imposing an additional constraint for the expected value of the outcome. This added constraint creates a substantial improvement in the consistency of the shape generated by competing distributions. As more constraints are imposed, the available space will continue to diminish, and some distributions will likely be eliminated for failing to simultaneously satisfy all constraints. When the experts have reached the limit of their ability to impose further constraints, the remaining step is to then choose the CDF that best mirrors their belief. This is the distribution / parameter set that will be recorded in the CRM.

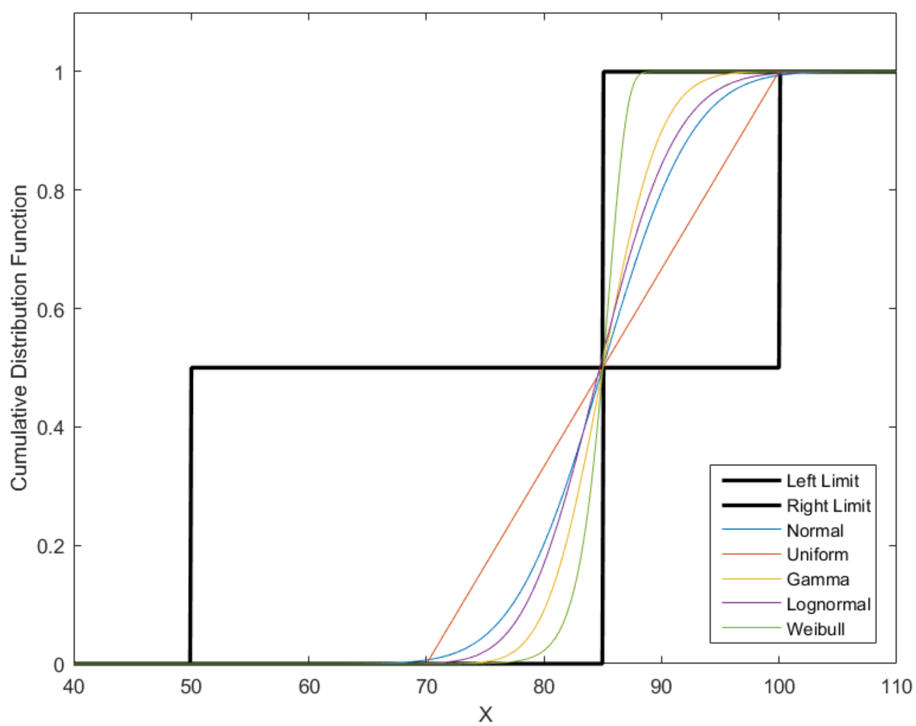


Figure 19: CDF Bounds for Minimum, Maximum and Expected Value Constraints

المنارة للاستشارات
www.manaraa.com

4.1.3 Linguistic Elicitation of Uncertainty

The process of eliciting uncertainty in terms of probability bounds requires a response from the Functional SME along the lines of the following: *"I expect the requirement to be X, but it will be no less than Y and no greater than Z"*. Posing questions in such a fashion is known as the direct method of elicitation. While this may seem to be the optimal approach, several authors assert that this has approach has historically been shown to be unreliable [11, 129]:

> This method elicits a direct estimate of the degree of belief of an expert on some issue. Despite its simple nature, this method might produce the worst results, especially from experts who are not familiar with the notion of probability

When prompted to explain the uncertainty or confidence in estimates, experts tend to use more descriptive language. Such a response would therefore be more likely to be structured as follows: *"It's likely that the requirement will be X, and highly improbable that it would be greater than Y or less than Z"*. Although these linguistic terms are somewhat fuzzy, they are meaningful. Lichtenstein and Newman conducted an extensive survey of the use of this kind of language to elicit expert opinions; a summary of which is provided in Table 5 [104]. The responses from subjects demonstrated considerable consistency in defining each term, but the range of quantitative responses paired with these definitions is significant. It can therefore be concluded that linguistic probabilities provide a useful start point, but some form of expert feedback is required to validate these probability bounds.

4.1.4 Maximum Entropy Formulations

The process described above can be summarized as follows:

1) Functional SME's use descriptive language to articulate their confidence in various estimates.

المنارة للاستشارات
www.manaraa.com

Table 5: Linguistic Interpretations of Probability [104]

Rank	Phrase	No. of Responses	Mean	Std. Dev.
1	Highly Probable	187	0.89	0.04
2	Very Probable	187	0.87	0.07
3	Quite Likely	188	0.79	0.10
4	Usually	187	0.77	0.13
5	Likely	188	0.72	0.11
6	Rather Likely	188	0.69	0.09
7	Somewhat Likely	187	0.59	0.18
8	Fair Chance	188	0.51	0.13
9	Uncertain	173	0.40	0.14
10	Possible	178	0.37	0.23
11	Rather Unlikely	187	0.21	0.10
12	Improbable	187	0.12	0.09
13	Very Unlikely	186	0.09	0.07
14	Rare	187	0.07	0.07

2) Probability bounds on the true CDF are derived from this description.

3) CDF's of known distributions are constructed to accommodate the corresponding constraints; experts then review the results.

4) Either a closed-form CDF is chosen and the process is complete, or the probability bounds are refined and the process repeats.

Unfortunately, a potential outcome of this iterative process is a failure to converge to a mutually agreed upon result. Multiple authors argue that this outcome, particularly in light of Arrow's Impossibility Theorem[1], is not only possible, but quite likely [11, 71, 129]. In these instances, the overwhelmingly preferred course of action is to leverage the Principle of Maximum Entropy (PME) to find an acceptable distribution.

The PME formulation is predicated on what Laplace referred to as the *principle of insufficient reason*. The principle of insufficient reason simply states that if one

[1]Arrow's Impossibility Theorem can be summarized as stating that each SME, acting as an individual, can exhibit rational behavior, but the decision-making of the group, taken as a whole, might appear utterly irrational [77].

www.manaraa.com

wishes to assign probabilities to an event and sees no reason for one outcome to occur more often than another, then the events should be assigned equal probabilities. The PME follows this line of thinking by enforcing the requirement that when information becomes available suggesting non-uniformity in outcomes, then the best distribution should be consistent with the knowledge provided without introducing any further assumptions [41]. Edwin Jaynes, the whose seminal work created this branch of probabilistic analysis, found that the notion of entropy in information theory afforded a direct measure of the amount of information expressed by a given distribution. Jaynes explains this concept as follows [93]:

> The great advance provided by information theory lies in the discovery that there is a unique, unambiguous criteria for the amount of uncertainty....[therefore], in making inferences on the basis of partial information we must use that probability distribution which has maximum entropy subject to whatever is known. This is the only unbiased assignment we can make; to use any other would amount to arbitrary assumptions of information which by hypothesis we do not have.

The definition of entropy for discrete and continuous variables is provided by Equations 3 and 4. A minimum to these functions, subject to constraints, can then be determined as the solution to the Lagrange multiplier problem shown in Equation 5 [41].

$$h(p) = -\sum_{i \leq 1} p_i log[p_i] \tag{3}$$

$$h(p) = -\int_I p(x) log[p(x)] dx \tag{4}$$

المنارة للاستشارات
www.manaraa.com

$$
\begin{aligned}
L(p_1, ..., p_n, \lambda_1, ...\lambda_n) = -\sum_{i\geq 1}^{n} p_i log[p_i] + \lambda_1(\sum_{i\geq 1}^{n} p_i - 1) + \sum_{j\geq 1}^{n} \lambda_j p_i \\
\frac{\partial L}{\partial p_1} = ... = \frac{\partial L}{\partial p_n} = \frac{\partial L}{\partial \lambda_1} = ... = \frac{\partial L}{\partial \lambda_n} = 0
\end{aligned}
\tag{5}
$$

A significant shortcoming of the PME approach is the types of constraints that can be incorporated into the above equations. Specifically, PME can accommodate constraints on the support (e.g. minimum between a and b) and moments of the unknown distribution (e.g. expectation, variance, skewness, and kurtosis). While experts can reasonably be expected to provide estimates for the supports and expectation, it is far less likely that they would be able to provide accurate estimates on the higher order moments. Consequently, part of the potential utility of the PME approach is likely to go unused. What is more problematic, however, is the fact that there is no well-formed method to incorporate CDF observations beyond the minimum, maximum and mean values. Thus, any observations used to generate CDF bounds beyond these data points must be sacrificed in order to apply this methodology [41, 138].

With that in mind, the PME approach can still be quite useful in certain scenarios. Table 6 provides a list of some of the common scenarios in which this method could be deemed sufficient.

4.1.5 Review

In review, time variation in requirements for OSA are intended to support an evolutionary development strategy for system capabilities. The long-term objectives guiding this directed evolution are codified in strategic road maps, but these road maps do not provide the type of hard requirements that could be incorporated into the design process. The true requirements, and the time-line along which they will be imposed, are ultimately governed by the outcome of parallel development programs (e.g. advanced FMV sensors). This dependent relationship introduces considerable

المنارة للاستشارات
www.manaraa.com

Table 6: Common Maximum Entropy Formulations [11]

Constraints	*Maximum Entropy Distribution*
$\int_a^b f_X(x)dx = 1$ Minimum Value $= a$ Maximum Value $= b$	Uniform
$\int_a^b f_X(x)dx = 1$ Expected Value $= \overline{X}$ Minimum Value $= a$ Maximum Value $= b$	Exponential
$\int_{-\infty}^{\infty} f_X(x)dx = 1$ Expected Value $= \overline{X}$ Variance $= \sigma^2$ Maximum/Maximum Values Unknown	Normal
$\int_a^b f_X(x)dx = 1$ Expected Value $= \overline{X}$ Variance $= \sigma^2$ Minimum Value $= a$ Maximum Value $= b$	Beta

uncertainty for the OSA platform intended to host these capabilities, and the necessary information to properly model this uncertainty is likely beyond the reach of the PM and their staff. These observations help explain why PMs are reluctant to embrace the OSA design philosophy - they are required to assume risk over which they have no control.

To resolve this challenge, this work posits the existence of a Functional SME with the responsibility to synthesize the various drivers of uncertainty into a single forecast. There are numerous methods to quantify uncertainty, but the context of the

المنارة للاستشارات
www.manaraa.com

problem dictates that a probabilistic representation is the most germane. Multiple authors acknowledge that best practices in translating expert forecasts into a probabilistic representations are based on establishing bounds for the CDF of the unknown RV. Closed form solutions can then be fit to these bounds using either existing distributions of the PME approach, though some iteration may be required to achieve an acceptable measure of agreement across a group of Functional SME's.

This process therefore requires the PM to gather additional information that is not currently considered in most existing design methods. The benefit to the program manager, however, is that following this process transfers the assumption of risk from the PM to the decision-makers and Functional SME's. Consequently, developing a CRM in this manner should not be considered as merely a "best practice", but a vital enabler for improving the overall implementation of OSA design principles. Thus, the linguistic bound and fit process described in this section is deemed to be an appropriate response to Research Question 1a.

With this in mind, the first step in the proposed methodology is to specify the purpose of an open design concept in the form of a CRM using the methods previously described. Figure 20 updates the process model with this observation.

4.2 Functional Analysis and Modular Partitioning

The methods identified in the previous section provide appropriate mechanisms to define requirements as an uncertain variation in threshold effectiveness over time. The intent of any open design is to efficiently satisfy these evolving requirements by upgrading components later in life, where uncertainty is reduced and component cost/performance are improved. These benefits cannot be realized, however, if the internal layout lacks sufficient modularity to prevent the design changes from propagating throughout the system. When this occurs, even small changes to the system can require substantial time and resources to implement. It is therefore imperative

www.manaraa.com

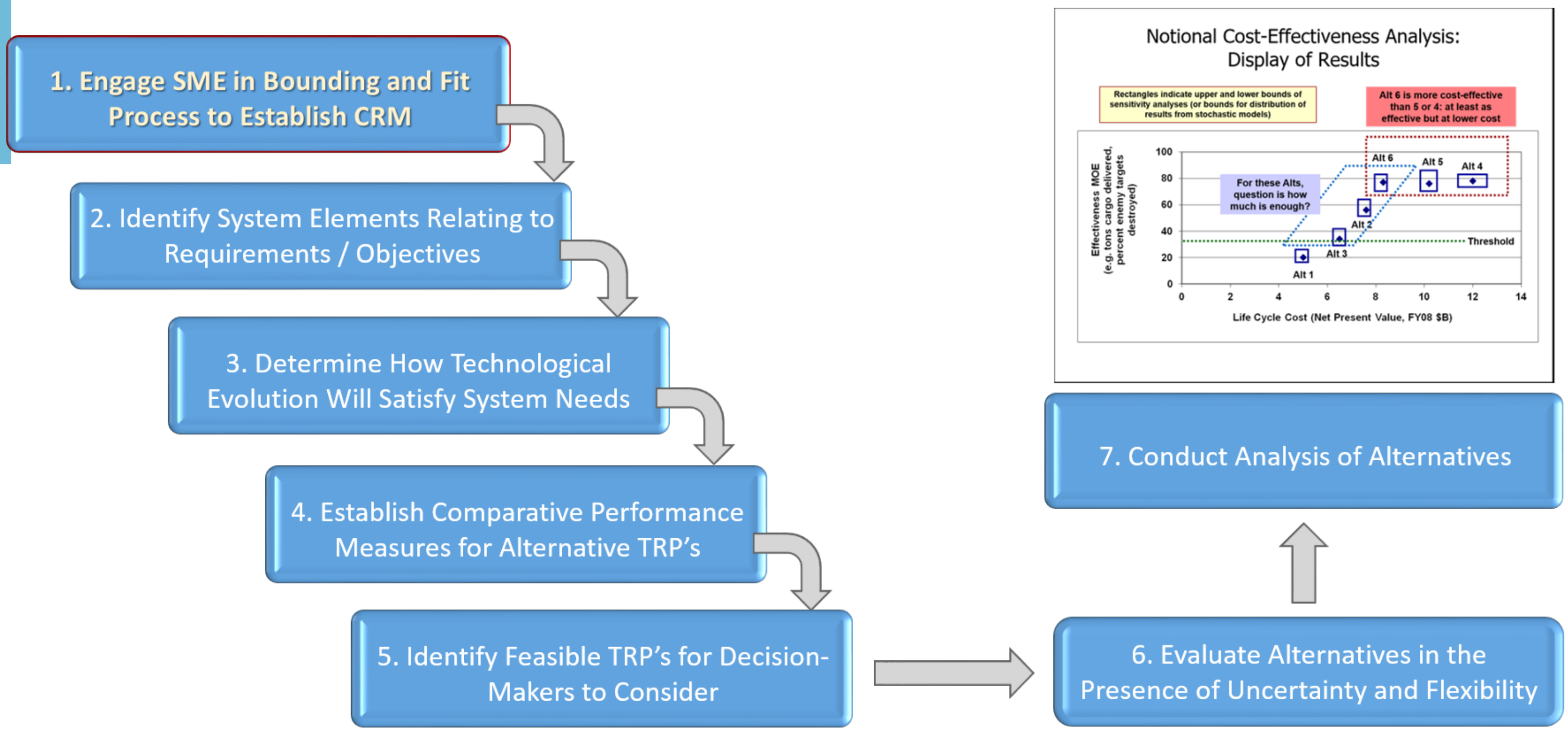


Figure 20: Methodology Update: Step One - Establish the Need via Capability Road Map

المنارة للاستشارات
www.manaraa.com

that designers identify a complementary set of modules to pair with the CRM requirements. Yet, it is not clear how much modularity is required for a successful OSA design, nor is it clear how systems functions should be organized into modules in order provide the greatest alignment with the CRM. This section will explore these issues.

4.2.1 Quantifying Modularity

There are numerous definitions of modularity in engineering design literature. Pahl defines modular products as "machines, assemblies, and components that fulfill various overall functions through the combination of distinct building blocks or modules" [137]. Allen and Carlson-Skalak define a module as a component or group that can be removed from the product non-destructively as a unit, which provides a unique basic function necessary for the product to operate [4]. Ulrich argues that modularity depends on two factors - similarity between the physical and functional architectures of designs and minimization of incidental interactions between physical components [174]. Marshall et al. describe modules as having the following characteristics [117]:

- They are co-operative subsystems that form products, manufacturing systems, etc.
- Functional interactions occur within, rather than between modules.
- They have one or more, well-defined functions that can be tested in isolation from the system and are a composite of components of the module.
- They are independent and self-contained, and can be combined and configured with other modules to achieve overall functions.

www.manaraa.com

There are clearly numerous competing definitions as to what constitutes a module and what it means to be "modular". There are, however, consistent themes underlying the volumes of literature surrounding modular design. A survey conducted by Gershenson et al. groups these themes into the following three categories: the independence of a module's components from external components, the similarity of components in a module with respect to their life-cycle process, and the absence of similarities to external components. The authors also provide a more apropos, albeit flippant, summary of these observations [69]:

> The only consensus in this review is that all believe a modular product is made up of modules, building blocks. The more components that fit into these modules, as opposed to lying around independently, the more modular a product is. The definition of modularity is therefore built upon the definition of modules.

Given the observations surrounding the definitions of modules and modularity, it comes as no surprise that there is a lack of academic consensus as to how one measures the extent of modularity in a design. However, there is some degree of consistency among competing measurements in that they should be determined from a system's Design Structure Matrix (DSM) [159]. The purpose of the DSM is to represent the hierarchical relationships and inter-dependencies between elements of the system. Note that the word "elements" is used deliberately here, as the DSM can be used to represent that relationships between system functions, physical components, development tasks, etc., depending on the context of the problem and the method being applied. To construct this matrix, one simply lists the various n elements present in the model and records the index of each entry. An empty n by n matrix, A, is then formed where the rows and columns represent the elements with the corresponding index. Finally, the elements of the DSM, A_{ij}, are populated with values indicating

المنارة للاستشارات
www.manaraa.com

either the existence of a relationship, or a numerical value quantifying the strength of the relationship. Values in the upper triangular region indicate a feed-forward of relationship (i.e. A_{ij} indicates that the j^{th} element receives an input from the i^{th} element), and the lower triangular portion is populated in the same manner to indicate a feed-back. Figure 21 provides an example application of the DSM approach to the task structure of a semiconductor development effort.

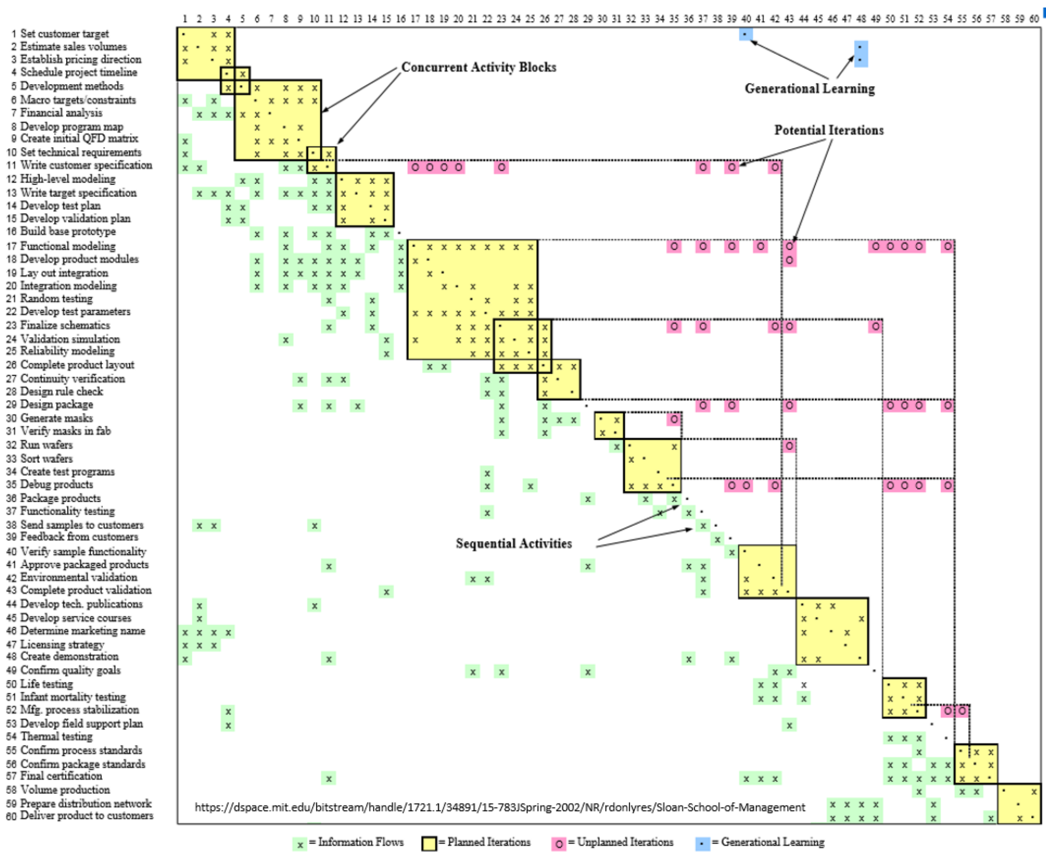


Figure 21: Application of DSM methods to Semiconductor Development [161]

The goal of DSM methods is to determine the optimal arrangement of rows and columns such that the diagonal of the matrix can be decomposed into sub-matrices containing the maximum amount of feed-forward/feed-back. These sub-matrices represent modules, and the remaining off-diagonal terms outside the sub-matrices and the on-diagonal overlap within sub-matrices represent interactions between modules. There are numerous competing methods to optimize the DSM and its constituent

المنارة للاستشارات
www.manaraa.com

modules. The majorities of these techniques, however, are distinguished only by the algorithm used to order the DSM. These nuances are not material to the present discussion; rather, the concern is with how a DSM organized in this manner can be used to quantify the modularity of a design.

4.2.1.1 Coupling Approach

One of the simplest mechanisms for quantifying modularity is the coupling approach developed by Lipson et al [106]. The coupling approach expresses the DSM as a real valued matrix, where off-diagonal terms quantify the degree of dependency on a relative scale (i.e. $A_{ij} \in [-1, 1] \; \forall \; i, j$). This matrix is reordered to form block matrices, and then reduced to a modular coupling matrix. The authors argue that any measure of modularity should be inversely proportional to the amount of coupling in the system. Under the theory of parsimony, the simplest metric to quantify this relationship is given by Equation 6, which is simply the sum of the magnitude of off-diagonal elements[2] across all N components divided by the total number of possible connections.

$$C_A = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} |A_{i,j}|}{N^2} \tag{6}$$

The authors note that their metric correlates well with with a system's ability to satisfy changing requirements. They concede, however, that there is no hard dividing line to indicate when a system transitions from "rigid" to "flexible"[3]. Rather, the metric, and by extension their interpretation of flexibility, is a relative property. In other words, when considering two competing designs, one can use this metric to determine which candidate is more modular or flexible, but it cannot determine how

[2]Note that by the authors' definition, $A_{ii} = 0, \; \forall \; i$.

[3]The authors reference the work of Saleh et al. when referring to the notion of flexibility. In this context, flexibility is defined as the ability of a design to satisfy changing requirements after the system has been fielded [148].

www.manaraa.com

much flexibility is sufficient, let alone optimal.

4.2.1.2 Network Approach

Sosa et al. also use a DSM approach to quantify modularity, though they approach the problem from a graph theoretic perspective [157]. In this formulation, the components are represented as nodes in the graphs, and the "connections" among these components are represented by edges connecting the corresponding nodes. In addition, the *degree* of a node is defined as the number of incident edges, the *length* of a path is the number of edges contained in the set, and the *geodesic* is the shortest possible path connecting two points. A graph is said to be *connected* if every pair of nodes is joined by a path, and a *bridge* is any edge whose removal would disconnect the graph. Finally, the authors' formulation requires that all edges of the graph are directed[4], and paired with an quantitative assessment of the strength of this relationship, x_{ij}. The DSM under this representation is then defined as the adjacency matrix for the network of components.

Applying this formulation provides numerous potential metrics for measuring the extent to which modularity is present in a design. The simplest measure, known as *Degree Modularity*, is given by Equation 7. This metric operates on the basic assumption the more components that affect or are affected by component i, the less modular that component should be. This property is expressed by the second term in the equation, which is simply the ratio of the actual degree of the node to the maximum possible degree - one in which a node is connected to every other node at the highest level of dependency, x_{max}. Finally, inverting the indicies of the summation in the numerator provide separate metrics for *in-degree* and *out-degree* modularity, but both metrics are structured such that increasing values reflect increasing modularity of the i^{th} component [157].

[4] A directed graph is one in which the edges have arrows indicating the direction of the corresponding relationship. This allows for asymmetric and symmetric relationships.

المنارة للاستشارات
www.manaraa.com

$$M_{degree,i} = 1 - \frac{\sum_{i \neq j}^{n} x_{ij}}{x_{max}(n-1)} \tag{7}$$

Although degree modularity captures how many other components are directly impacted by a given component, it does not consider any of the indirect relationships. The authors argue that this consideration is best quantified by evaluating how distant the component is from other components in the network. They therefore define *distance modularity* as a measure proportional to the sum of the geodesics for the i^{th} component, $d(i,j)$, with all other components. This metric, provided in Equation 8, therefore depends on the direction of dependencies, but not the strength of those dependencies. As with degree modularity, a distinction can be made between *in-distance* and *out-distance* modularity by inverting the indicies for the geodesic function. The intent of both metrics is the same; higher values imply greater isolation, and therefore greater modularity [157].

$$M_{dist,i} = \frac{\sum_{i \neq j}^{n} d(i,j)}{n(n-1)} \tag{8}$$

The final metric for modularity proposed by the authors is built on the assumption that components that lie on the most geodesics are those bridging the most components, and are therefore the least modular. The authors argue that this assumption is appropriate in the product domain when design dependencies propagate through the minimum umber of parts (i.e., the geodesic). According to this logic, an appropriate measurement can be defined by considering the ratio of all geodesics connecting components a and b that contain the i^{th} component, $nd_{a,b}(i)$, to the total number of geodesics, nd_{ab}. This comparison yields a measure of how much i^{th} component contributes to the bridge connecting the corresponding nodes. An aggregate measure of *bridge modularity* can then be determined summing over all pairs of components and normalizing the result with the total number of possible paths. This expression

المنارة للاستشارات
www.manaraa.com

is provided in Equation 9. Note that, as with the previous metrics, higher values indicate less connectivity, and thus greater modularity [157].

$$M_{bridge,i} = 1 - \frac{1}{(n-1)(n-2)} \sum_{i \neq a, i \neq b, a \neq b} \frac{nd_{ab}(i)}{nd_{ab}} \tag{9}$$

The network approach advocated by the authors appears to be a substantial improvement over the coupling-based method previously established. This approach is particularly appealing due to the fact that the various metrics have an intuitive, physical significance. Yet, the authors acknowledge that there is no readily apparent method to aggregate the various metrics into a single, holistic measure of system modularity. Moreover, the degree modularity metrics share a similar challenge to the coupling method in that there is no consistent method to measure the strength of coupling between dependent elements in the network.

4.2.1.3 Binary Representations of Connectivity

Binary representations of the DSM seek to quantify the degree of modularity of a product based solely on its internal connectivity structure. As such, the DSM is populated with zeroes along the diagonal, and the off-diagonal terms are set to unity if two components are connected either physically, or through the transmission of power or information. One of the most widely cited methods in this field is the Singular Value method developed by Holtta et al. [80]. The indicators of modularity in this methodology are defined by the eigenvalues of a Singular Value Decomposition on the binary DSM. These parameters are determined through the solution of Equation 10, which is provided as Equation 11 below:

$$DSM = U \cdot \Sigma_{DSM} \cdot V^T \tag{10}$$

www.manaraa.com

$$\Sigma_{DSM} = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_N \end{bmatrix} \tag{11}$$

In Equation 11, the singular values σ_1,σ_2,....σ_N are ordered in decreasing magnitude, and N is the number of components in the system. The authors investigated the physical significance of these values by performing the singular value analysis on three canonical system architectures: fully integrated, fully modular, and bus-modular[5]. A depiction of the architectures and the corresponding eigen-structure are shown, respectively, in Figures 22 and 23. As depicted, there is an obvious difference in the singular value decay structure for the various archetypes. The integral system has one large singular value followed by N-1 smaller ones, the bus-modular system has two mid-range singular values with the remaining values at or near zero, and the fully modular system has a slow, gradual decay. The authors offer the following explanation for these observations [80]:

> The information content to describe the connectivity of the system is different for modular versus integral systems. The modular design, for example, requires more information to describe completely, relative to the bus-modular system. In other words, all singular values.....must be retained for a complete description of the modular system.

These observations led to the conclusion that highly integrated systems can be identified by the extent to which the important information necessary to describe the system is concentrated in a few, highly connected components. These systems exhibit

[5]Bus modularity is a common modular configuration, and is commonly defined as any device with two or more interfaces that accepts any combination of components from a set with standard interfaces [161].

المنارة للاستشارات
www.manaraa.com

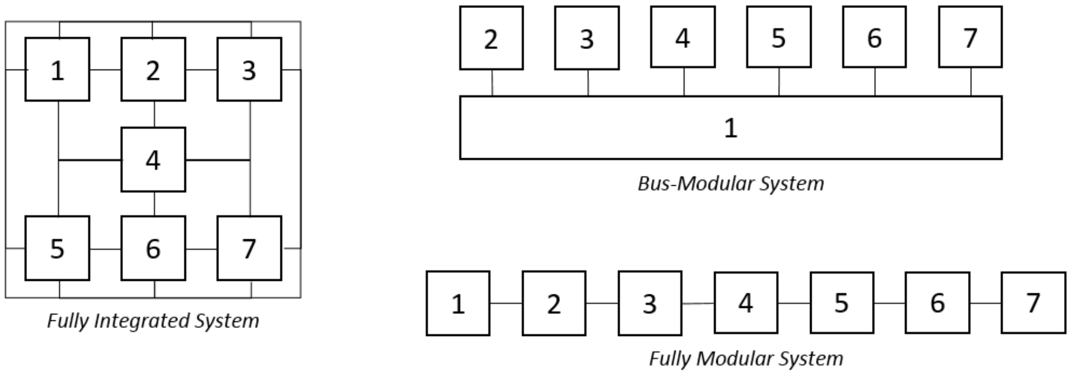


Figure 22: System Idealization Archetypes [80]

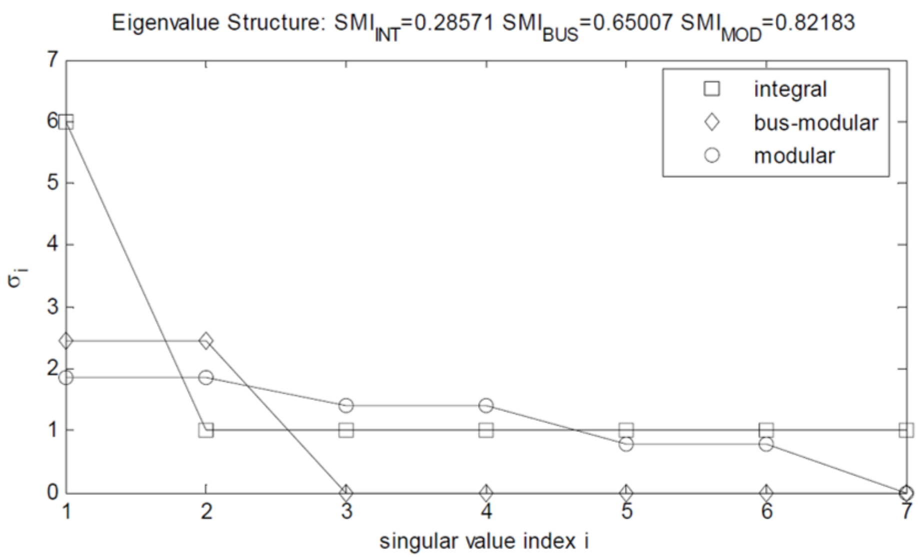


Figure 23: Eigen-Structure for System Idealization Archetypes [80]

a rapid drop-off in the magnitude of their singular values relative to modular systems, where the information is more widely distributed. Thus, the authors argue that an appropriate measure of modularity is the average, weighted rate of decay across the sorted singular values in the system. This metric is referred to as the Singular-value Modularity Index (SMI), which is provided in Equation 12 below.

$$SMI(\Sigma_{DSM}) = 1 - \frac{1}{N \cdot \sigma_1} \sum_{i=1}^{N-1} \sigma_i(\sigma_i - \sigma_{i+1}) \qquad (12)$$

This index is bounded between zero and one, with increasing values indicating a

www.manaraa.com

higher degree of modularity. For the example problem provided in Figure 23, the authors determined the following: the integrated design has an SMI score of 0.29, the bus-modular design has a score of 0.65, and the fully modular design has a score of 0.82. These trends tend to support the original hypothesis. It should be noted, however, that these are idealized cases, and there are no real world observations to provide benchmark SMI scores. More data will need to be collected before the SMI can be used to determine how a system performs along the real-world spectrum of integrated vs. modular designs. In addition, the computational burden of optimizing the SMI score of product architectures may become problemmatic as the the number of components, and thus the size of the DSM, increases [80].

In summary, modular design has been an area of intense academic interest in the engineering community. This interest has led to numerous definitions of what it means to be "modular", as well as an impressive number of metrics to quantify the extent of modularity in a system design. Yet, the current challenge lies in relating these idealized metrics to real world systems in order to establish benchmarks to serve as necessary and/or sufficient conditions for a product to be deemed "modular". Gershenson rather succinctly explains this barrier to real world application [69]:

> Most of the effort is put into proposing new ideas as opposed to testing existing or proposed hypotheses. Notable is the lack of comparison among the implications of varying definitions of modularity and the measurement and validation of proposed benefits and costs.

4.2.2 Heuristic Methods

Barriers exist between theoretical measures of modularity and their significance with respect to real world systems which limit their potential utility in a generic OSA methodology. There are, however, heuristic strategies that can be applied to develop modular partitioning schemes that do not rely on such measures. As the name implies,

المنارة للاستشارات
www.manaraa.com

these methods specify *a priori* heuristics about what a good module should look like, and then search the design for combinations of functions and/or components that meet these criteria. Heuristic methods therefore make no observations about the relative costs and benefits of a given partitioning scheme; they simply seek to identify potentially useful modules. It is then left to the designer to determine if elementary modules should be integrated into the system architecture, segregated as a module, or combined with other elementary modules to form a larger module.

There are numerous heuristics in this field, but the most popular come from the seminal work of Stone and Wood [161]. The authors idealize the complexities of modular design as a binary decision for the designer; products can have an integrated architecture, where functional elements map to a single or very small number of physical components, or a modular architecture, where physical components have a one-to-one correspondence with the product's functional model. They propose the methodology depicted in Figure 24 to explore the latter alternative.

Steps one, four, and five of this approach are common engineering practices, and are therefore not the focus of the method. The starting point for Step Two is a functional decomposition of the design, which requires that the overall function of the product be hierarchically decomposed into progressively smaller, more detailed functions. The lowest level of decomposition then provides an elementary set of functions used to describe the flow of energy, material, and signals flowing through the product[6] [137]. This representation is referred to as a "Black Box" model in the sense that there is no physical component at this stage to dictate how these functions are to be executed. These basis functions are then linked together to define the operations conducted on each flow from its entrance into the product until its exit. Combining the individual functional chains provides the aggregate functional model

[6]Potentially useful enablers in this process are the functional bases developed by various authors to aid in such a decomposition [107, 160].

المنارة للاستشارات
www.manaraa.com

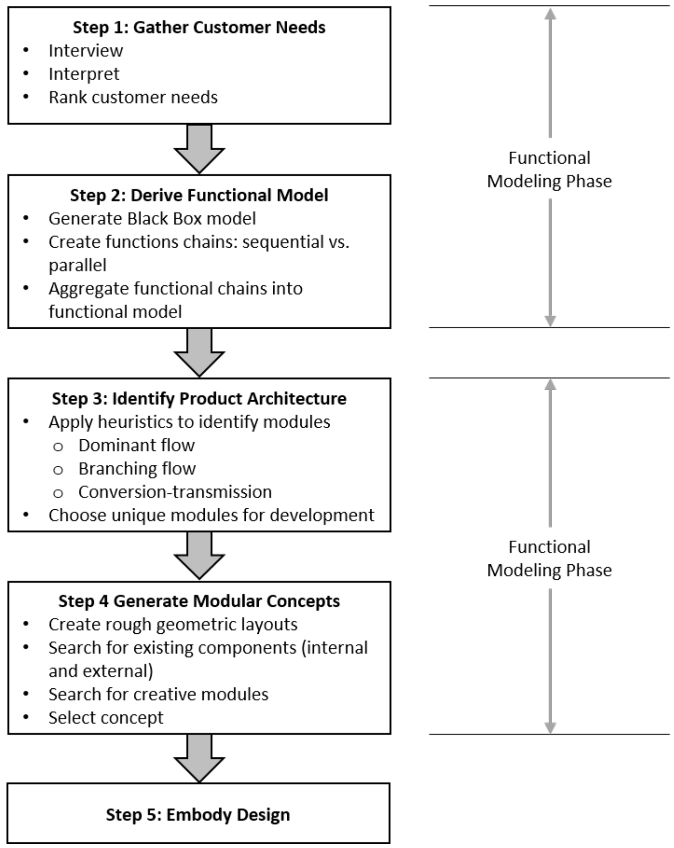


Figure 24: Heuristic Identification of Modular Partitions [161]

necessary for the next step in the methodology.

The focus of Step Three is to determine where modular partitions should be introduced to the aggregate model. The authors argue that three heuristics can be applied to manage this process. The first heuristic, **Dominant Flow**, defines modules as the set of sub-functions governing any non-branching flow through the model. Other flows, to include the traced flow, that cross this boundary are the interfaces indicating an interaction with another module. **Branching Flow** defines modules in terms of parallel function chains. By definition, this requires all modules to interface with the product at the flow's branch point, which lends itself well to identifying instances of bus-modularity. **Conversion-Transmission** is the final heuristic, which, as the name suggests, defines modules as functional sets that follow the flow of material or energy up to, and including, its conversion to another form of material or

المنارة للاستشارات
www.manaraa.com

energy. Interfaces of a conversion-transmission module are identified using the same techniques of both branching and dominant flow modules [161].

The authors note that the three different heuristics will likely identify overlapping modules, as well as modules which are sub-sets or super-sets of other heuristics. As previously mentioned, however, the authors do not provide a cost-benefit relationship to determine which set of modules is ideal. The general guidance provided for these situations is to chose those modules with the lowest number of sub-functions in order to be consistent with the philosophy that modules should be easily identifiable with a particular function. In addition, there is an immediate problem with the scaling of this method to systems with greater complexity. To articulate this point, consider the functional decomposition of a screwdriver provided in Figure 25. The authors identified 12 potential modules in this functional decomposition, but it is quite clear that identifying these modules will become increasingly difficult as the complexity, and therefore the size of the functional model, of the underlying product grows.

In review, heuristic methods do not seek to determine the optimal configuration of a modular design. Rather, they identify a set of basis components/modules, and leave the decisions of what to do with a given module to the designer. Compared to the quantitative DSM models, this approach is appealing in that there is a logical, physically intuitive justification for the identification of modules. However, the challenges associated with scaling this approach to problems of greater complexity imply that this method alone is likely to be insufficient for the OSA problem.

4.2.3 Design Rules

Design Rules is an alternative approach to managing modular design considerations that lies at the intersection of heuristic methods and DSM-based quantification of modularity. The methodology is part of a greater theory put forward by Baldwin and Clark to explain the rapid pace of design evolution in computer engineering [15].

المنارة للاستشارات
www.manaraa.com

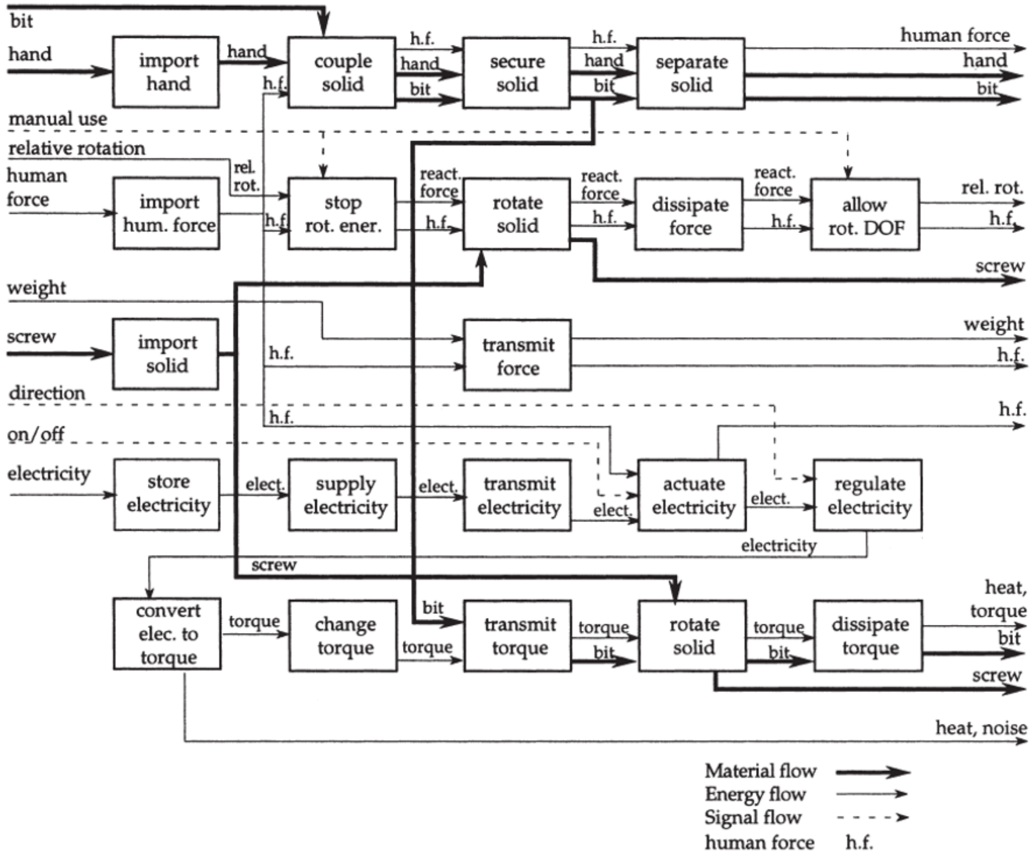


Figure 25: Heuristic Identification of Modular Partitions [161]

The central thesis of their work is that the use of modularity in computer hardware and software is what made this miracle possible. To that end, they propose a system-agnostic methodology to quantitatively measure the added value of progressively decomposing a system into smaller modules. This work has proven to be extremely popular in academic literature, and numerous authors have extended their work into fields ranging from organizational hierarchies [27] to the development of therapeutic protocols in Applied Psychology [37].

The DSM in the Design Rules approach is slightly different from the binary, coupling and dependence structures previously described. Here, the rows and columns of the matrix represent DV or development tasks, as opposed to design variables. The authors justify this representation through the argument that a fundamental

www.manaraa.com

isomorphism exists between the design structure and task structure during system development. Thus, if one solves the task structure problem, which may be conceptually simpler, then it is possible to find the equivalent design structure. In addition, entries on the lower triangular portion signify a hierarchical relationship (e.g. parameter x_1 calls parameters $x_2 \ldots x_j$ into being), as opposed to a feed-back relationship. Feed-forward and feed-back relationships are considered to be instances of "interdependency", which are recorded in the upper triangular region using either a binary representation or a qualitative intensity scale (e.g change in x_i makes a change in x_j desirable). The authors go on to state that reordering methods for the DSM are useful, but typically insufficient. They give the example shown in Figure 26 to support this point, along with the following commentary.

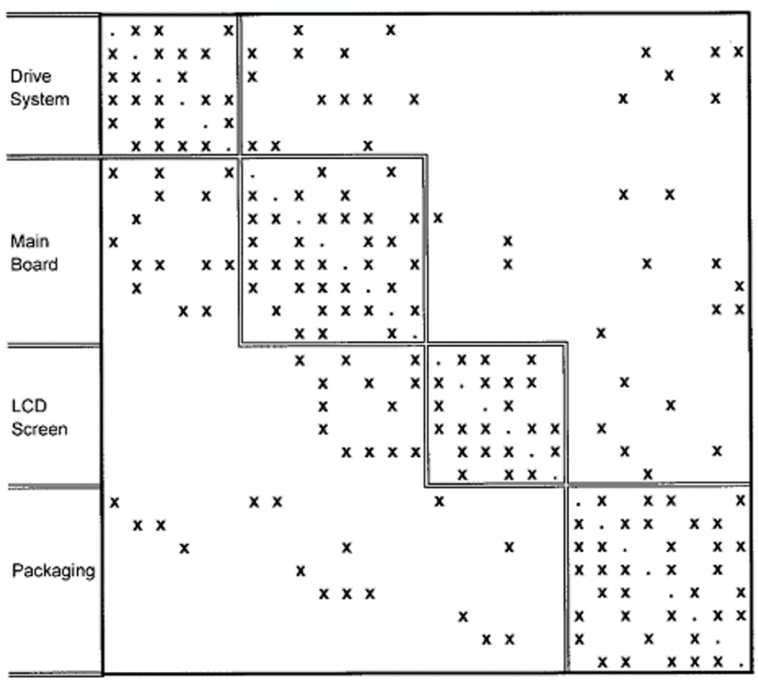


Figure 26: DSM Reordering without Design Rules [15]

How does the work of design proceed with a design and task structure such as this? The tasks do not fall into a natural sequence. If one starts at the top of the matrix and works through the tasks in a fixed order,

المنارة للاستشارات
www.manaraa.com

cyclical interdependencies…quickly drive one "back" to reconsider previous decisions [15]

To remedy this shortcoming, the authors posit the existence of an "architect". The role of the architect is to consider which of the off-diagonal terms should be kept within the design process, and which entries should be fixed *a priori*. This type of decision benefits the design team by reducing complexity, but it comes at the cost of restricting a potentially useful region of the design space. For example, one of the entries in Figure 21 is the decision to include a graphics processor. If the architect knows this decision will not have a major impact on the design, then the decision can be made in advance. Decisions of this nature are referred to as Design Rules.

In addition, not all design rules are created equal. The hierarchical nature of systems means that some decisions taken at a lower level will have little to no impact on other system elements. The authors account for this notion of "information hiding" by classifying Design Rules according to whether they have a system wide impact, or are contained within a sub-matrix of the sorted DSM. Figure 27 depicts the impact of imposing hierarchical Design Rules on the original DSM of Figure 26. Note that the Design Rules are consolidated in the leftmost columns, with the global Design Rules at the top.

This process produces a block hierarchical design and task structure, but these blocks are not necessarily the optimal modular configuration. In many instances, it is preferable to further decompose these blocks into smaller modules, which may, in turn, benefit from further decomposition. Yet, it is not clear how those sub-modules should be identified, nor is there an apparent indicator as to when this decomposition should stop. To better manage this exploration phase of decomposition, the authors define complete set of six heuristic operators that have been proven to be successful in previous modular design efforts. Alternative partitions can then be developed by repeatedly applying the operators in different combinations. These operators are

المنارة للاستشارات
www.manaraa.com

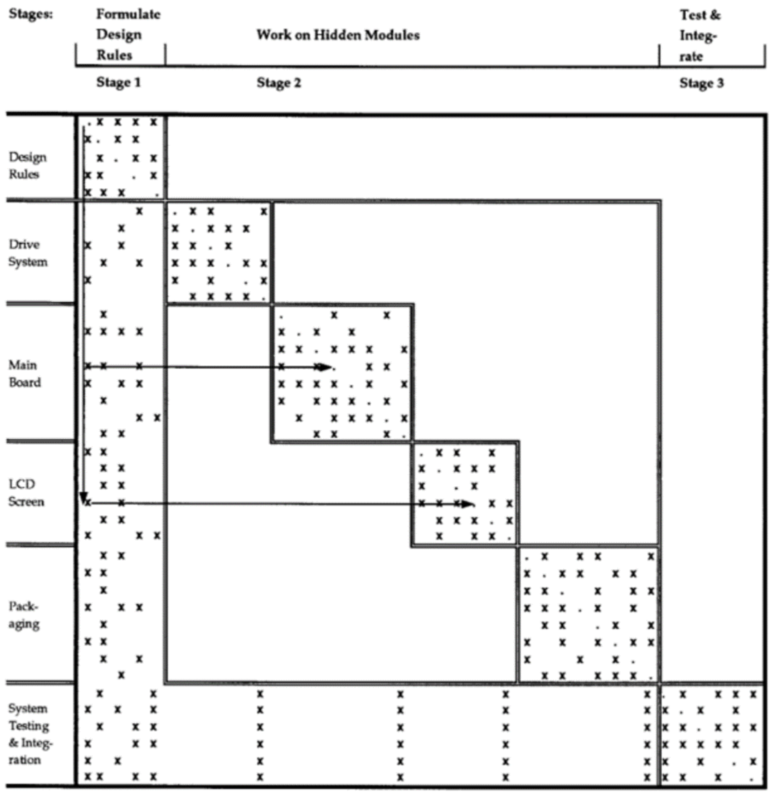


Figure 27: DSM Reordering with Design Rules [15]

listed below:

- *Splitting* a design (and its tasks) into modules
- *Substituting* one modular design for another
- *Augmenting* the system by adding a new module
- *Excluding* a module from the system
- *Inverting* to create new design rules
- *Porting* a module to another system

The modular operators are useful for generating alternative partitions, but they do not inherently provide guidance as to which alternative is preferable. To resolve this, the authors argue that a set of modules represents a portfolio of design options,

www.manaraa.com

analogous to those found in financial markets. They therefore apply financial valuation techniques to understand the market value of individual design/decomposition decisions. A comprehensive treatment of the concepts underlying financial options beyond their immediate application to this method is beyond the scope of this dissertation[7]. However, it is worth noting that the basic premises of "options thinking" appear in numerous modular design applications. This observation will be explored in greater detail in later sections.

With that in mind, application of the financial options paradigm to the modular design problem is formulated as follows. Assume that the value provided by forming the i^{th} module, X_i, can be modeled as a normal distribution with a mean of zero and variance proportional to the number of tasks embedded within the module. This implies that $X_i \sim \mathrm{N}(0, \sigma^2 \alpha_i N)$, where N is the total number of tasks within the sub-matrix, and α_i is the percent of those tasks allocated to the i^{th} module. For simplicity, the random variables are converted to a standard Normal distribution using the Gaussian transformation provided in Equation 13. When the time comes to implement this module, designers will only proceed with the proposed module if its perceived value is greater than zero. Consequently, the expected value of the i^{th} module is given by Equation 14, and it follows that the overall value across all options is given by Equation 15 [15].

$$z_\alpha = \frac{X_\alpha}{\sigma \left(\alpha N\right)^{1/2}} \tag{13}$$

$$E\left[X_i | X_i > 0\right] = \sigma \left(\alpha N\right)^{1/2} \int\limits_0^\infty z_\alpha dz_\alpha = 0.3989 \sigma n^{1/2} \tag{14}$$

$$V_\alpha = 0.3989 \sigma N^{1/2} \sum_{i=1}^{n} \alpha_i^{1/2} \tag{15}$$

[7]The interested reader is referred to the author's original text for a more detailed explanation [15].

المنارة للاستشارات
www.manaraa.com

The equations listed above provide a baseline concept of value, but this value corresponds to one partitioning scheme. In practical applications, however, it is quite likely the designers will formulate parallel design efforts to experiment with different modular configurations. A substantial benefit of modularity comes from the decoupling of these experimental outcomes, which allows the design team to select the "best of breed" across k alternatives. This distribution of the best of k designs is well known in statistics as the distribution of the "maximum order statistic of a sample size k". The maximum order statistic is defined by Equation 16, where N and n represent, respectively, the standard normal CDF and PDF. Substituting this relationship into the previous value formulation simplifies to the formulation in Equation 17, the total value of a design process with j modules and k parallel experiments per module.

$$Q(k) = k \int_0^\infty z\left[\mathrm{N}(z)\right]^{k-1} \mathrm{n}(z) dz \tag{16}$$

$$V(j,k) = \sigma \left(Nj\right)^{1/2} Q(k) \tag{17}$$

Of course, modularity is not free. Experiments are costly to run, and their results must be tested and integrated into the greater architecture. The Net Option Value (NOV) is therefore the difference between value gained from introducing modularity and the cost incurred to produce that value. Equation 18 provides the final form of NOV advocated by the authors. Here, c_j is the design cost per module, c_k is cost per experiment, and $T(j,k)$ is the cost of testing j modules with k experiments per module. As the goal of a commercial venture is to maximize profits, it follows that the optimal modular layout can be determined by finding the modular operator chain and the number of experiments per module maximizing Equation 18[8].

[8]The authors provide this formulation for the foundational splitting/substitution operators. Variations of this formulation exist for the remaining operators, but the information presented here is sufficient to demonstrate the underlying valuation concepts.

المنارة للاستشارات
www.manaraa.com

$$NOV(j,k) = \sigma\,(Nj)^{(1/2)}\,Q(k) - c_j j - c_k k - T(j,k) \tag{18}$$

In summary, this methodology has a three-phase approach to managing modular design considerations. First, Design Rules and reordering algorithms are used to divide the DSM into a set of isolated block matrices. Second, a set of heuristics are applied to the independent sub-matrices to identify alternative partitioning schemes. Finally, valuation methods are applied at the lowest level of decomposition to determine the value of the corresponding set of modules. This approach is appealing when compared to the purely heuristics-based approach because the presence of Design Rules resolves the challenge of scalability in more complex products. However, the specifics of the valuation method are problematic when compared to the acquisitions OSA. For example, acquisition programs are interested in both cost and effectiveness, and it is not clear how these competing objectives can be reduced to a single criterion to replace the objective of maximizing commercial profit. Moreover, even if such a criterion existed, it is not clear how the numerous drivers of uncertainty can be reduced to a single normal distribution. Thus, Design Rules provides useful insight, but it is unlikely to independently resolve Research Question 1b.

4.2.4 Axiomatic Design

The final method under consideration for identifying modular partitions is the theory of Axiomatic Design proposed by Suh [163]. Axiomatic design is an overarching framework to formalize the theory of design, and is therefore intended to be applicable, albeit with slight variations, to both open and closed system architectures. The foundation of this framework is the segregated design spaces depicted in Figure 28. Suh argues that all designs originate in what he refers to as the customer domain, which is characterized by the customer's attributes (CA), the attributes that the customer desires in the final product. The role of the engineer is to translate these

المنارة للاستشارات
www.manaraa.com

desires into Functional Requirements (FR) and constraints that must be satisfied for the design to be successful. In order to translate function to form, the FR must ultimately map to Design Parameters (DP) in the physical domain. Finally, to produce the specified product in terms of the DP, a production effort must be established in the process domain. This process is characterized by Process Variables (PV).

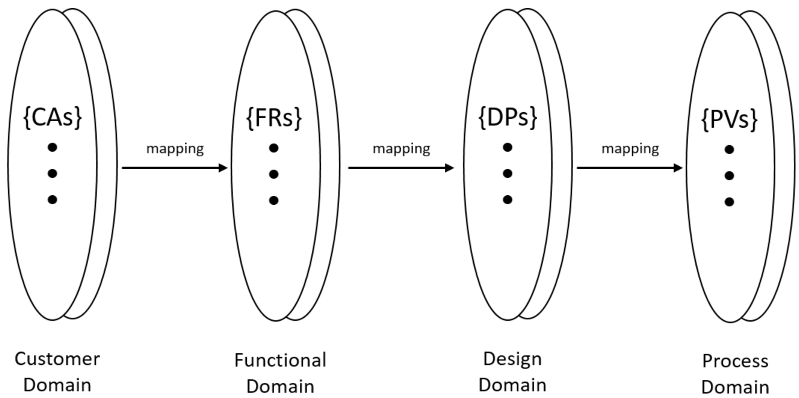


Figure 28: Axiomatic Design Spaces [163]

The DSM in Axiomatic Design is unique among DSM methods in that the rows and columns do not represent the same quantity. Rather, the purpose of the DSM is to provide a mapping between the functional domains shown in Figure 28. This relationship, the "design equation" in author's vernacular, is given in Equation 19 for the mapping between the functional and design domain. Consequently, rows of this DSM, A, would correspond to the functional domain, columns would correspond to the design domain, and individual entries would reflect the change in an FR caused by a change in the corresponding DP. Individual matrix entries are constants for a linear design, whereas for nonlinear designs they are a function of the DV. Note that this is a second deviation from other DSM based methods, where the entries are either binary or qualitative measures of interdependencies.

$$\{FRs\} = [A]\,\{DPs\} \tag{19}$$

A second purpose the DSM in Axiomatic Design is to enforce the *Independence*

المنارة للاستشارات
www.manaraa.com

Axiom. The Independence Axiom states that when there are two or more functional requirements, the design architecture must be such that each of the functional requirements can be satisfied without affecting the other functional requirements. In the context of the DSM, this axiom is satisfied if, and only if, the matrix is either diagonal or triangular. Diagonal matrices imply that that each FR can be satisfied independently by a single DP. Systems with this property are referred to as *uncoupled designs*. If the DSM is triangular, known as *decoupled designs*, then the independence of FR can only be guaranteed if the DV are varied in the proper order [162].

If a DSM satisfies the Independence Axiom, then Axiomatic Design provides a provision to identify the optimal partitioning of system modules. A module in this context is defined as the row of a design matrix that yields an FR when it is provided with the input of the corresponding DP. Equation 20 provides a simple design example to illustrate how this is applied in practice.

$$\begin{Bmatrix} FR1 \\ FR2 \end{Bmatrix} = \begin{bmatrix} a & 0 \\ b & c \end{bmatrix} \begin{Bmatrix} DP1 \\ DP2 \end{Bmatrix} \tag{20}$$

Let M1 and M2 be the modules corresponding to the elements of the DSM that yield FR1 and FR2 as a univariate function of DP1 and DP2 respectively. By definition, M1 and M2 must satisfy the relationship given in Equation 21. The solution to this system of equations therefore provides the definitions of M1 and M2 according to Axiomatic Design.

$$\begin{aligned} FR1 = \;& aDP1 + 0DP2 = & M1 \cdot DP1 \\ FR2 = \;& bDP1 + cDP2 = & M2 \cdot DP2 \end{aligned} \implies \begin{aligned} M1 = \;& aDP1 \\ M2 = \;& b\left(\tfrac{DP1}{DP2}\right) + c \end{aligned} \tag{21}$$

As shown, the Independence Axiom provides a unique modularization scheme for a system as a function of its DPs. This is a unique property compared to the previous methods, as it does not allow for the consideration of alternative modular descriptions. While this might be considered a substantial drawback of the method, it should be

المنارة للاستشارات
www.manaraa.com

noted that one of the barriers of success in OSA implementation is the conflict of interest created by the potential for vendor lock. Thus, it may actually be beneficial to embrace a consistent, codified process for determining a modular structure that leads to single decomposition. The greater challenge arises when individual components contribute to two or more DPs. In this case, it is possible, even guaranteed in the case of uncoupled designs, that the component will appear in two or more modules. It is not clear how such a conflict could be resolved in a consistent manner.

4.2.5 Conclusion

In review, there are numerous methods to determine a modular partitioning scheme for a design. Quantitative methods seek to assign a measure of "goodness" to a given scheme, but the relative nature these methods complicates their use in real world analyses. In other words, these methods are well-suited to determine if one scheme is more modular than another, but there is little research to ensure a consistent threshold for sufficient, let alone optimal, modularity is achieved. Heuristic methods bypass this problem by pursuing a bottom-up approach. These methods yield a large set of elementary modules and leave the designer to determine whether each module should be integrated, aggregated, or segregated. This ensures that no potentially useful modules are missed, but the approach becomes increasingly difficult to implement when the size and complexity of the system increase. Design Rules follow a more centric path by using DSM methods to isolate specific aspects of the system, at which point heuristics can be applied to develop alternative partitioning schemes. Yet, the quantitative metrics used to identify the "optimal" partition schemes are intended to accommodate commercial ventures, where the sole objective is to maximize profit. As such, this approach is unlikely to gain acceptance in the acquisitions context, where there are potentially multiple MoEs contributing to a more abstract concept

المنارة للاستشارات
www.manaraa.com

of "value". Finally, Axiomatic Design provides a consistent, traceable method to determine the best partitioning scheme for a design. Unfortunately, this approach will likely provide conflicting results if a real world design fails to conform to the idealized assumptions underlying the development of methodology (i.e. each component provides a single DP in an uncoupled design).

The conclusion then is that there is no silver bullet to identify necessary / sufficient conditions for modularity in an OSA design. This review does, however, lend considerable insight into the best practices that should be applied during development. Namely, the approach should be:

- **Consistent:** Provides a repeatable process applicable to a wide variety of systems
- **Scaleable:** Possesses a linear scaling between the level of effort required and the size/complexity of the design
- **Complete:** Ensures that all useful modules are identified

To that end, this work advocates a two-phase approach as a baseline methodology. In the first phase, Design Rules are applied to divide the modular partitioning space into reasonable sub-spaces. In the second, heuristics, or, in highly ideal circumstances, Axiomatic methods can be used to identify a modular basis worthy of further consideration. Such an approach would be consistent, scaleable, and complete, and is therefore deemed to be an appropriate resolution to Research Question 1b. It should be noted, however, that there is no academic basis to define a precise method for establishing the Design Rules or heuristics that should be applied. These decisions must be governed by the context of the problem at hand.

Finally, the basis elements identified by this process can be mapped to the CRM in order to determine which portions of the system architecture contribute to the the stated objectives and requirements. This is the stated requirement for the second

المنارة للاستشارات
www.manaraa.com

step of the proposed methodology. Figure 29 updates the process model with this observation.

المنارة للاستشارات
www.manaraa.com

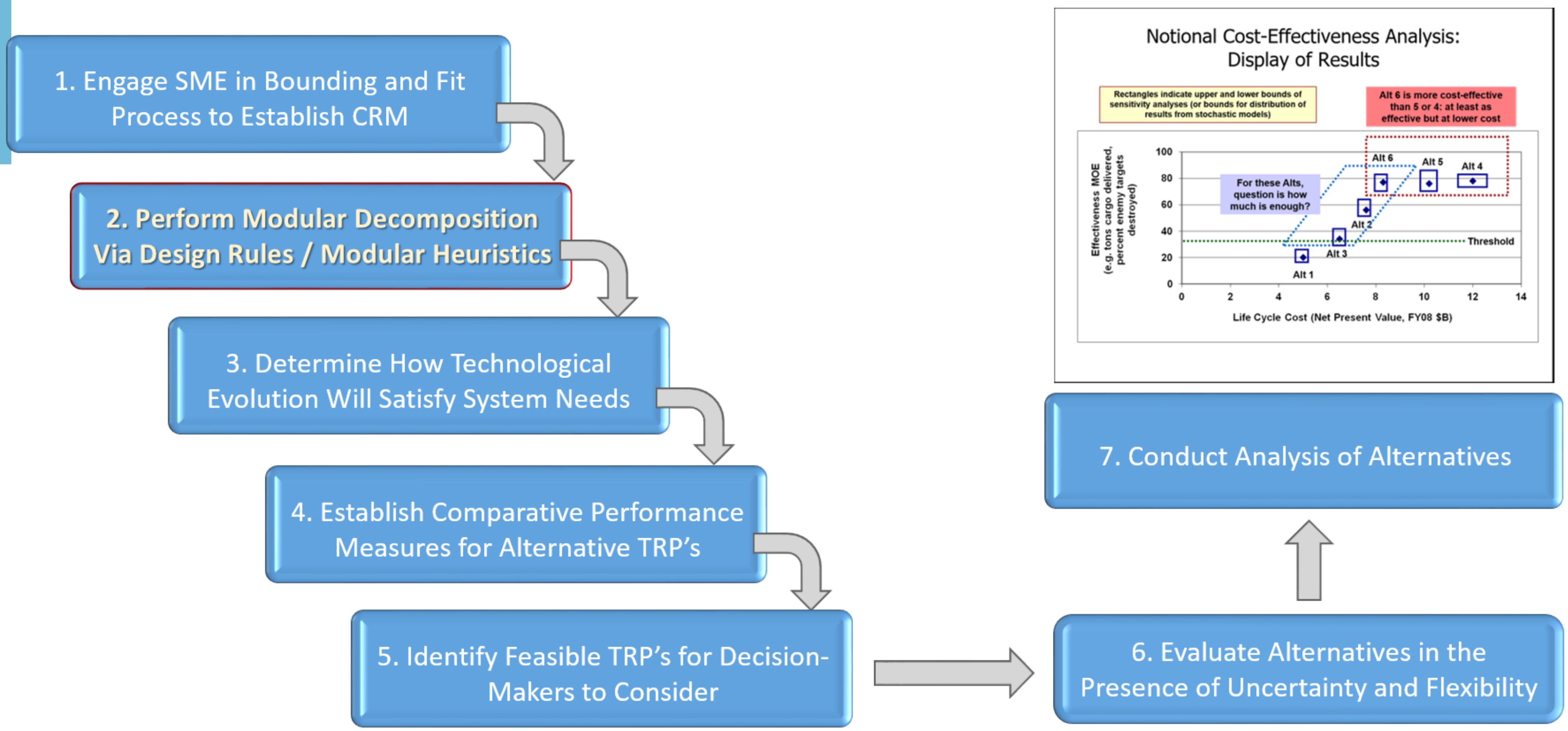


Figure 29: Methodology Update: Step Two - System Decomposition via Design Rules and Modular Heuristics

المنارة للاستشارات
www.manaraa.com

CHAPTER V

GENERATING TECHNOLOGY REFRESH PLANS

The previous section developed methods to ensure that time variation in system requirements is documented in a CRM, and to identify modules that can be upgraded to satisfy these requirements. Decision-makers will not, however, accept that an OSA will be capable of efficiently satisfying these evolving requirements on faith. Open systems must undergo the same cost-effectiveness scrutiny to which all other acquisition programs are subjected. This requires, among other things, a TRP to be presented alongside the initial system concept to project when the various components will be upgraded, what the resulting system effectiveness will be relative to projected thresholds, and the TLCC. It is not immediately clear how this TRP can be constructed, however, as the government has little to no ability to direct the process of technological evolution in commercial markets. The purpose of this section is to develop a method to overcome these obstacles.

5.1 Multi-Object Decision Making Techniques

5.1.1 Terminology and Notation

A useful starting point in the search for methods to develop and select a TRP is to take a closer look at the techniques used to evaluate the cost and effectiveness of traditional systems. This will require a more refined and formalized vernacular than what is commonly used in the acquisitions literature considered in Chapter Two. To that end, this work will follow the common engineering convention formalized by Daskilewicz, where complex systems are represented by a vector of scalar quantities that completely define the system under consideration [47]. This implies that any parameters not contained within this set can be assumed to be either fixed, or irrelevant to the analysis

المنارة للاستشارات
www.manaraa.com

at hand. The properties within this set that can be independently determined are referred to as *Design Variables (DV)*, and the properties dependent on the DV are referred to as *Response Variables (RV)*. The functional relationships mapping DV to RV are typically accomplished though computational analysis, which will henceforth be referred to as M&S[1].

Let n denote the number of design variables, and let x_i denote the particular value of the i-th DV. Each DV is further assumed to be contained within a corresponding set of bounding constraints, c_i, which can be either continuous (e.g. closed set [1,2]) or discrete (e.g. integer values 1 or 2). A design in which all DV lie within the prescribed bounds is known as a *feasible design*, and the sub-space of all feasible designs within $\Re^n$ is defined as the *Design Space*, D. An n-tuple of DV that defines a given design will be represented by a vector $\mathbf{x}$, which allows D to be formally defined as $D = \{x \in \Re^n | x_i \in c_i \text{ for } i \ = \ 1, 2, ...n\}$. In addition, let k be the number of RV under consideration, and define y_j be the particular value of the j-th RV. Similar to the DV notation, a k-tuple of RV is represented as a vector, $\mathbf{y}$. Thus, an instantiation of a design analysis is mathematically idealized as $\mathbf{y} = f(\mathbf{x})$, where is f represents the unknown functional mapping determined through M&S.

$$f : \begin{cases} \Re^n & \rightarrow & \Re^k \\ \mathbf{x} & \Rightarrow & \mathbf{y} \end{cases}$$

The *Objective Space*, O, is the image of the design space under the M&S transformation, i.e. $O = \left\{\mathbf{y} \in \Re^k \mid \mathbf{y} = f\left(\mathbf{x}\right), \ \mathbf{x} \in D\right\}$. The essence of the classical engineering design problem is to determine the point within O that provides the greatest value to the designer. It is assumed in this analysis that each dimension of O, i.e. each RV, has a corresponding objective to be either minimized or maximized. Thus,

[1]Note that the definitions of DV and RV are equivalent to the definitions of MoP and MoE provided in Chapter Two. The more formal vernacular is applied in this section to prevent confusion and simplify notation.

www.manaraa.com

if two designs have identical RV in all but one dimension, then the design with the greater RV would be preferred if the objective is to maximize, and the opposing design would be preferred is the objective is to minimize. The core challenge in this paradigm, however, is to determine the "best" design in the presence of multiple, competing objectives.

5.1.2 Partial Ordering - Pareto Optimality

The most general method to manage design decisions is to simply assume that the designer is unable to make any judgment on the preference structure relating different RV. A given design is considered to "dominate" another design if, and only if, every RV is superior to its counterpart in their respective objective (i.e. minimize or maximize). For example, if RV are intended to be minimized, then the design $\mathbf{y}$ is said to dominate the design $\widehat{\mathbf{y}}$ if $\mathbf{y} \neq \widehat{\mathbf{y}}$ and $\mathbf{y_j} \leq \widehat{\mathbf{y}}_\mathbf{j}\ \forall\ j = 1, 2...k$. If this is true, then one would always select the dominant design over the dominated design, thereby excluding the dominated design from further consideration. Consequently, this scenario does not permit the identification of a "best" design. Rather, the goal is to determine the set of non-dominated designs with the collective property that no RV can be improved without a penalty in at least one other desirable attribute. These designs are said to be *Pareto Efficient* and the space they occupy within O is referred to as either the *Pareto Frontier* or, equivalently, as the *Trade Space* [2]. The intent of this Pareto analysis is to then present the available options to the appropriate decision-makers in order to allow them to impose their preferences *a posteriori* [47]. Figure 30 depicts an example of the trade space available for a design characterized by two RV, both of which are intended to be minimized.

While conceptually simple, determining the exact shape of the Pareto frontier often proves to be challenging. The nature of this challenge stems from the fact that

[2] Trade Space is synonymous with the Pareto Frontier, and the two will be assumed to be equivalent in this work.

المنارة للاستشارات
www.manaraa.com

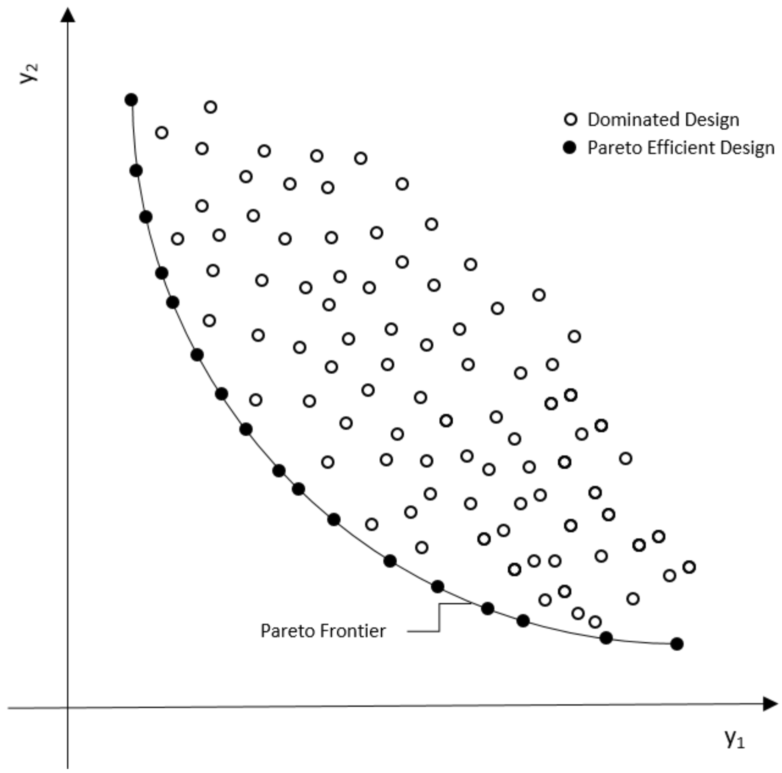


Figure 30: Notional Pareto Frontier

the M&S mapping $\mathbf{x}$ from the Design Space, D, to $\mathbf{y}$ in the Objective Space, O, is often highly non-linear. Therefore, whereas D possesses a simple hyper-cube structure, there is no such convenient description for O, let alone the sub-space defining the Pareto frontier within O. Identifying the structure of this Trade Space is an area of considerable interest in the field of Engineering Design and Optimization, and two classes of methods have proven to be quite effective at managing this task: iterative optimization and evolutionary algorithms [47].

5.1.2.1 Iterative Optimization

One of the most common methods to identify the Pareto frontier is the Aggregate Objective Function (AOF) approach. As the name implies, AOF methods combine the various RV into an overall objective function of the form $y_{Agg} = f(w_1 \cdot g(y_1),$ $...w_j \cdot g(y_j))$. Here, g_k represents a scaling/normalization function of the k-th RV,

المنارة للاستشارات
www.manaraa.com

w_k represents the relative importance of the k-th RV to the overall design objective[3]. This form allows a given point on the Pareto Frontier to be determined by solving a single univariate optimization problem: *minimize* y_{Agg}, *subject to* $\mathbf{x} \in D$. As depicted in Figure 31, this process can be iterated for different combinations of weights to identify different points along the frontier [116].

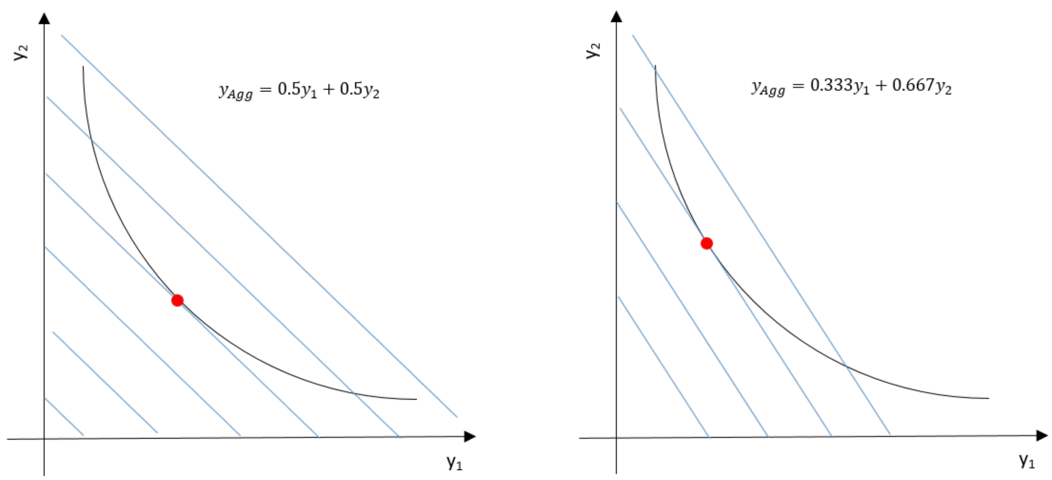


Figure 31: Notional AOF Method Application

It is important to note that the weighted sum technique is unable to sample points on a concave portion of the frontier, and those points identified on convex portions tend to have irregular spacing [116]. These qualities make the weighted sum approach somewhat unreliable for the generic Pareto exploration problem. Several authors have introduced alternative functional forms for the AOF that seek to remedy these shortcomings [115, 124]. Messac et al., however, have shown that the curvature of the AOF must be greater than the curvature of the concave potions of the frontier in order to accurately map these spaces [127, 125]. This observation has led to greater acceptance of the Chebyshev norm, which has infinite curvature, as a consistent method of identifying non-dominated solutions across an arbitrary frontier. Unfortunately, this form still suffers from an inability to evenly distribute sampled points along the

[3]By convention, $\sum_{k=1}^{j} w_k = 1$.

www.manaraa.com

frontier [47].

An alternative to the AOF approach is to convert all but one of the RV to either equality or inequality constraints, and then allow the remaining RV to serve as a univariate objective function[4]. In this formulation, the $k-1$ equality/inequality constraints are evaluated at a grid of points ranging from a minimum threshold to a maximum value in their respective dimension of O. The sub-optimization problem defined in Equation 22 is then solved at each point on the constraint set to identify a corresponding Pareto efficient point. Figure 32 demonstrates the practical application of this concept on the familiar two RV problem used in previous discussions. Note that this method is similar to the AOF approach in the sense that the "weight" parameters, α_j are contained within the optimization constraints as opposed the objective function itself [101, 101].

$$\begin{aligned} &min && y_j \\ &subject\ to && \mathbf{x} \\ & && y_1 \leq \alpha_1, ..., y_{j-1} \leq \alpha_{j-1}, y_{j+1} \leq \alpha_{j+1}, ..., y_k \leq \alpha_k \end{aligned} \tag{22}$$

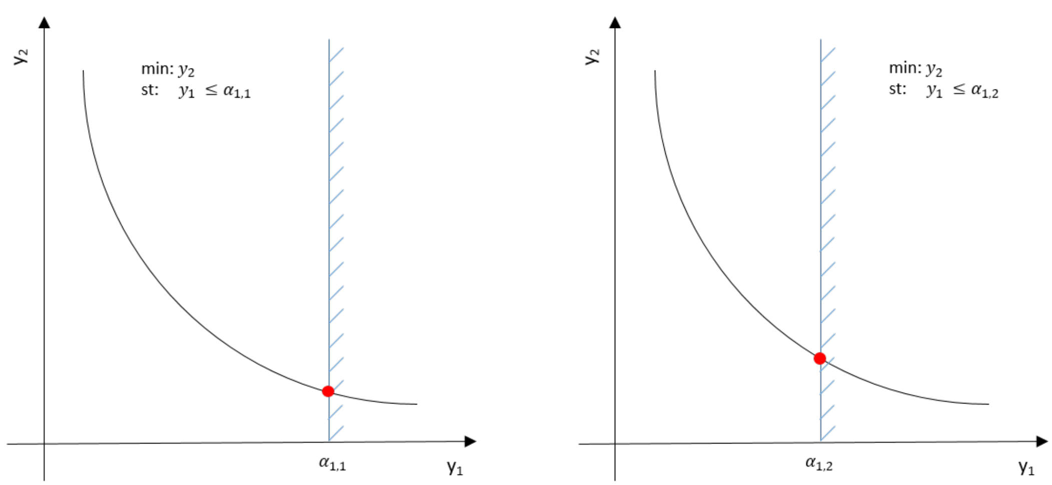


Figure 32: Notional Epsilon Constraint Method Application

[4]This method is referred to as the ϵ constraint method for inequality constraints and the Method of Proper Equality Constraints for equality constraints.

المنارة للاستشارات
www.manaraa.com

The final class of iterative techniques is somewhat similar to the ϵ constraint method, in that they form a series of sub-problems by perturbing a set of equality or inequality constraints. The key difference in this class of methods is that the constraint set is developed through a more sophisticated method that is intended to reduce the number of failed cases. One of the most widely studied methods in the class is the Normal Boundary Intersection (NBI) approach[5] [46]. NBI begins by using an appropriate single objective optimizer to find the k points that optimize each RV individually, resulting in the vector set $\{\mathbf{y}^*\}$. An evenly distributed simplex-grid of "basepoints" is then formed in the Objective Space, O, by taking weighted sums of the individual optima: $y_j^0 = w_1 y_1^* + w_2 y_2^* + ...w_2 y_2^*$. Next, the *Utopia* and *Anti-Utopia* points in O, μ^+ and μ^-, are identified by taking the best and worst attributes, respectively, found in $\{\mathbf{y}^*\}$: $\mu_i^+ = \inf\left\{y_{ji}^*\right\}$ and $\mu_i^- = \sup\left\{y_{ji}^*\right\}$ $\forall j = 1, 2, ..., k$. The normalized vector difference between μ^+ and μ^- then provides the search direction for each of the basepoints, $\overrightarrow{n}$. Each sub-problem in the exploration algorithm then consists of starting at a given base point, and then maximizing the distance along $\overrightarrow{n}$ under the constraint that $\mathbf{x} \in D$. This sub-problem is expressed in Equation 23 and a visual depiction is provided in Figure 33.

$$\begin{aligned} &max && t \\ &subject\ to && y_j^0 + t\overrightarrow{n} = f(\mathbf{x}) - \mu^+ \\ & && \mathbf{x} \in D \end{aligned} \tag{23}$$

5.1.2.2 Evolutionary Algorithms

The class of iterative optimization methods previously established share the common characteristic of attempting to sample the Pareto frontier by solving single-objective

[5] The Normalized Normal Constraint (NNC) method is a more sophisticated version of NBI which uses inequality constraints and a different mechanism to determine base points. Both methods function on the same principles, but the simplicity of NBI makes it more amenable to the current discussion. As such, NNC is not explicitly considered here [124, 126].

المنارة للاستشارات
www.manaraa.com

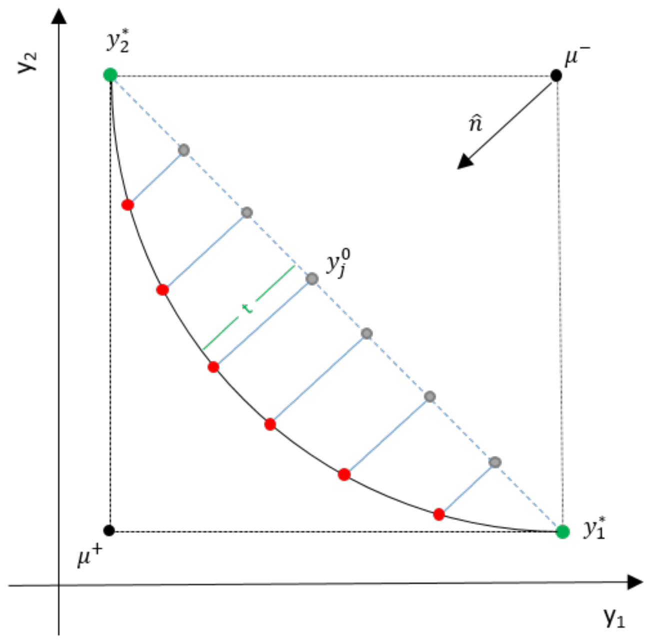


Figure 33: Notional Normal Boundary Intersection Method Application

sub-problems. Multi-Objective Evolutionary Algorithms (MOEA) take a fundamentally different approach to this challenge. Methods in this class operate on the concept that the problem can be dealt with through a single optimization executed on a population of points. The goal of this optimization is to migrate these points from their arbitrary start points to the frontier, and then maximize the separation of these points in order to ensure a uniform sampling of the trade space. Common algorithms in this class of methods include the Niched Pareto Genetic Algorithm[82], the strength Pareto Evolutionary Algorithm[179], and the Non-Dominated Sorting Algorithm - II (NSGA-II)[142]. The NSGA - II algorithm will find use in the methodology under development, and will therefore warrant closer scrutiny.

As previously mentioned, the goal of Pareto exploration is to sample points on the Pareto frontier, and to ensure that those points are as uniformly distributed as possible. NSGA - II directly integrates these concepts into its fitness function in the form of two quantities:

- Non-Domination Level: A measure of the relative dominance of a given point

المنارة للاستشارات
www.manaraa.com

with respect to the remainder of the population

- Crowding Distance: A measure of the distance between a given point and its nearest neighbors

Non-domination levels generalize the concept of dominance to an integer scale indicating how dominated points are within a population. To determine these values, a given population is mapped from D to O, and a Pareto filter is used to determine which points within the population are Pareto efficient. This subset of dominant points is assigned a dominance level of zero and are removed from the population. The process is repeated as many times as are necessary to assign all points a dominance score, with the dominance level increasing by one every iteration. Thus, lower dominance scores indicate that a point is closer to the true frontier, and a preference for lower scores helps drive the initial population towards that boundary.

Just as non-domination levels generalize the concept of dominance, crowding distance generalizes the concept of uniformity in the distribution of points along the Pareto frontier. This measure is only applied to points with the same non-dominance level. To determine these scores, the points within the current non-dominance level are sorted in ascending value according to their values in the first RV, y_1. The first and last entries within the set are assigned an infinite crowding distance to ensure that they automatically pass to the next generation. Distance scores for the remaining points are found by taking the average distance of the two nearest neighbors (i.e. directly above and below the current index in the ordered set). This process is then repeated across the remaining RV, $y_2, ..., y_k$, and the total score is determined by summing across the scores for individual attributes. In contrast to non-domination scores, greater values of crowding distance indicate greater isolation, and a preference for larger scores therefore drives points to maximize their distribution within their non-dominance level.

NSGA - II integrates these concepts into the classical Genetic Algorithm (GA)

المنارة للاستشارات
www.manaraa.com

framework. As with most Genetic Algorithms, NSGA - II begins by creating a randomly distributed population of n points in D, which are then mapped to O. The following generation (i.e. the children of the first generation) are formed through the standard crossover and mutation operators, and are evaluated in the same manner as the previous generation. The parents and children are combined into a single population of $2n$ points, and into their respective non-domination levels. Next, the algorithm determines the maximum non-domination level of the combined population such that fewer than n points have a non-domination level less than this threshold. This threshold is referred to as F_{limit}. The m points with a non-domination level less than F_{limit} automatically move to the next generation, and the points in levels below F_{limit} are excluded from further consideration. By definition, there are at least $n - m$ points remaining in F_{limit}. The crowding distance is calculated for each of these points, and the resulting set is ordered according to those values. Finally, the $n - m$ points with the greatest distance are added to the set continuing to the next generation, and the remainder are discarded. This process is then iterated until converged.

Reflecting on this algorithm, it is clear that the early generations will likely have many levels of non-dominance. Consequently, the non-dominance level will be the the most significant parameter in the fitness function, which equates to driving the population toward the frontier. Once that population reaches the frontier, however, the number of dominance levels drops precipitously, marking the second phase of the algorithm. In this phase, the crowding distance becomes the most significant parameter, which forces the population to spread along the frontier as evenly as possible. Note that this directly satisfies the two requirements for Pareto exploration algorithms.

In review, NBI[6] and the NSGA - II algorithm are the most commonly used methods for exploring the Trade Space available for a given design. Though both options

[6]Note that NNC, the more sophisticated sibling of NBI, is the more commonly used approach.

www.manaraa.com

perform well, there are implementation challenges to each. In particular, NBI assumes that both D and O are continuous spaces, and therefore has difficulty dealing with discrete parameters. To illustrate the significance of this limitation, consider an aircraft design problem that permits either one or two engines. In its native form, NBI would be willing to consider 1.5 engines as a viable solution, though it is clear that such a solution has no physical significance. NSGA - II, on the other hand, requires significantly more numerical function calls, and thus significantly greater time to execute. Further, there is no exact method for specifying the optimal convergence criteria, which implies that the algorithm may need to be run multiple times under different criteria to identify the true frontier. The context of the problem must therefore be considered when determining which algorithm is appropriate to the task at hand.

5.1.3 Implications of Time Variation in Performance

In light of the new vernacular and notation, it is clear that the cost-effectiveness methodology described in Chapter Two is a form of Pareto analysis. Yet, there is a rather obvious complication in any attempt to create a Pareto representation of open architectures - performance varies over time. The logical consequence of this observation is that the Pareto frontier will, under the assumption of improving technology, inexorably expand in the direction of the utopia point. In other words, when a technology refresh occurs, the available technology will possess some combination of improved performance attributes and/or reduced cost. Infusing this new technology in the greater system architecture will move the system to a state that was not feasible during the initial development phase. Thus, whereas closed architectures are evaluated in a static trade space, OSA must be evaluated in a dynamic trade space.

In addition, recall that a TRP defines when Technology Insertions (TI) will occur, and what type of technology will be infused. In the context of a dynamic trade space,

المنارة للاستشارات
www.manaraa.com

choosing when to upgrade is equivalent to selecting which frontier the system will move to, and selecting a given technology at that time is equivalent to identifying the point on the frontier terminating this movement. Connecting the points of this movement is therefore analogous to a literal development path that the system will follow. This leads to a new idea that will be central to this work - *Technology Refresh Plans can be idealized as a path through the trade space.* Figure 34 depicts a notional example of the implications these observations for an instance in which it is desirable to maximize two RV's. Note that this particular TRP consists of two TI's occurring at times t_2 and t_3, with component attributes indicated by the red dots.

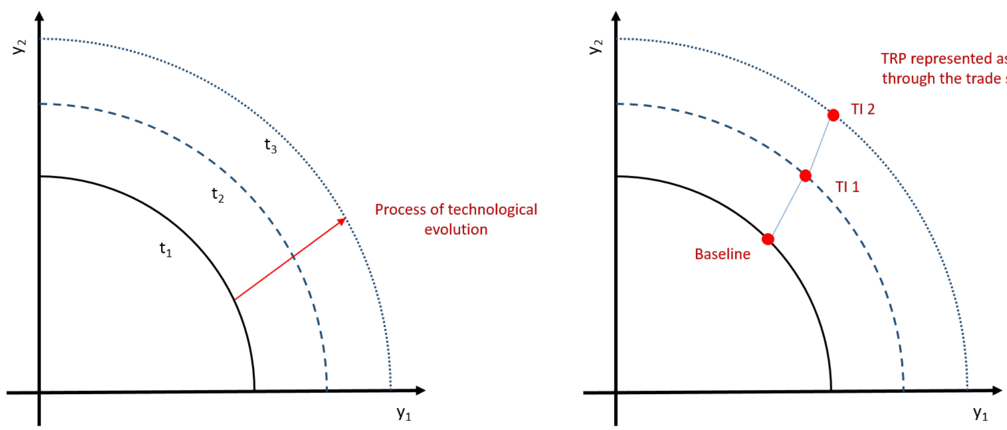


Figure 34: Idealization of TRP's as Paths Through the Trade Space

There is a second consequence to this formulation. Infusing advanced technology into a system can be idealized as a transformation of the system's design vector provided by Equation 24 below. In this equation, x_{t_0} is the baseline system, Δx_t is the marginal difference between the baseline components and the technology infused at time t, and $\widehat{x}_t$ is the new design vector at time. The formulation provided at the beginning of this chapter would therefore require that Δx_t propagate through M&S to create a corresponding improvement in the response vector, $\widehat{y}_t$. This relationship is provided by Equation 25.

www.manaraa.com

$$\widehat{x_t} = x_{t_0} + \Delta x_t \tag{24}$$

$$\widehat{y_t} = y_{t_0} + \Delta y_t = f\left(x_{t_0} + \Delta x_t\right) \tag{25}$$

For simplicity, assume that the two RV's depicted in Figure 34 combine to create a single concept of "system performance"[7]. If this is true, then the notional TRP would create variations in system performance over time. This work will refer to a total time history of performance variations as a *Performance Profile*. Figure 35 illustrates the basic concept of relating a TRP to a performance profile under the M&S transformation.

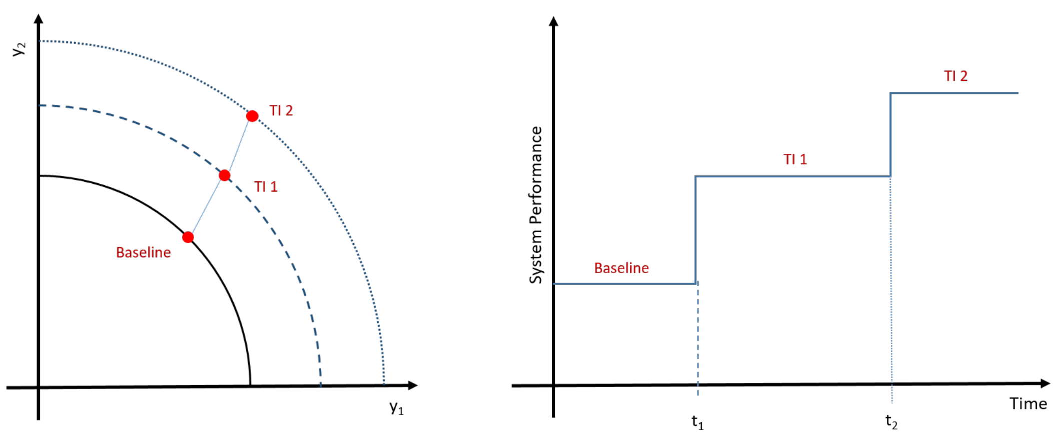


Figure 35: Performance Profile Created By a Notional TRP

Performance profiles can serve a useful purpose. It was noted at the beginning of the chapter that an important dimension of a TRP is the ability to demonstrate how future upgrades can be leveraged to meet evolving requirements. Consider again the notional CRM depicted in Figure 16, and define this time history of uncertain requirements as the system's *Requirement Profile*. Figure 36 juxtaposes the notional requirement profile generated by the system's CRM with the notional performance

[7]This assumption will be removed shortly, but it is advantageous at the present for visualization purposes.

www.manaraa.com

profile provided by the given TRP. It should be noted that uncertainty in both profiles must be considered, but this topic will discussed in greater detail in the next chapter. For now, it suffices to observe that the TRP satisfies evolving constraints if its performance profile is greater than the requirement profile for all time, t.

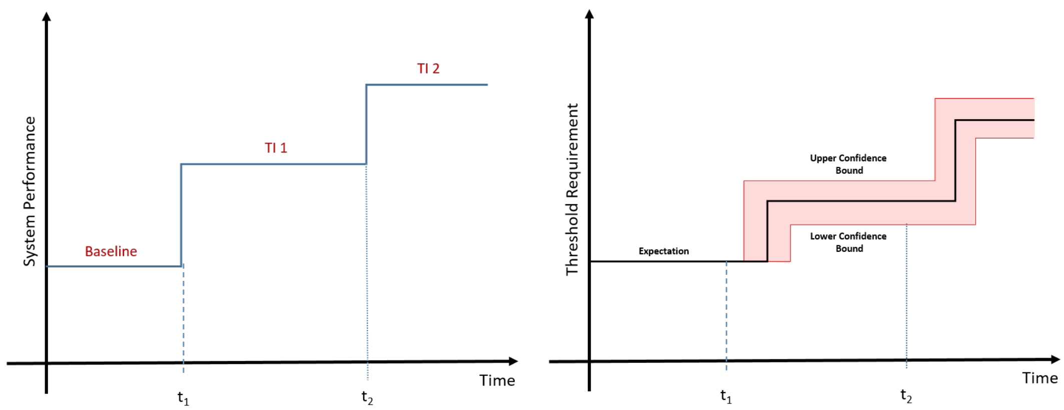


Figure 36: Juxtaposition of Performance and Requirement Profiles

In addition, this representation now makes it possible to define what it means for an OSA to *efficiently* satisfy requirements. In order to ensure that a system is operationally useful over its entire lifecycle, closed architectures define present-day requirements as those that are necessary to ensure a high degree of confidence that a system will satisfy its terminal requirements. This approach equates to working backwards from the extreme point on the performance profile, which results in the scenario depicted in right portion of Figure 37. To reiterate points made previously, satisfying these exacting requirements with present-day technology is believed to typically require custom-made, highly integrated components. This, in turn, leads to designs with some combination of higher weight, volume, power, cost, etc. Moreover, systems designed for terminal requirements may likely have a glut of excess capacity early in life that the system does not need and/or cannot use. In other words, no added "value" is received from this excess. Finally, the difference in efficiency between open and closed systems conveyed by Figure 37 is likely understated. The reason for this understatement is that uncertainty in the timing and magnitude of

www.manaraa.com

requirement increases is believed to diminish with time. Delaying the decision to infuse technology would therefore allow engineers to make more informed decisions as to how technology can be leveraged to meet the stakeholder's requirements.

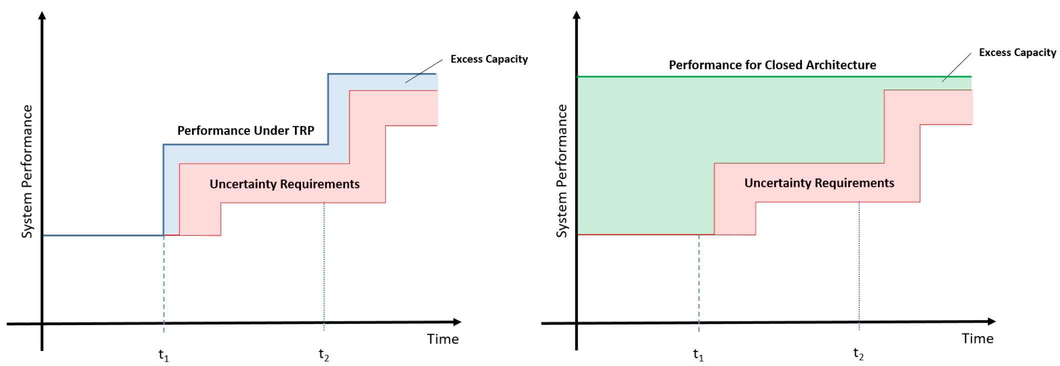


Figure 37: Excess System Capacity in Open and Closed Architectures

While this representation is informative, it also introduces new problems. In particular, it is somewhat self-evident that the development of a performance profile requires knowledge of how the State Of The Art (SOTA) in commercial components will evolve over time. The government has little influence over the direction of this evolution, and it is therefore unclear how analysts can consistently generate the forecasts necessary to make intelligent decisions regarding the construction of TRP's.

Another, and no less substantial, complication with this approach is that performance is no longer a vector of scalar attributes; rather, it is a set of profiles depicting how these attributes vary over time. It is unclear how MODM techniques, to include the concept of Pareto efficiency, can be applied to this formulation. This, in turn, is problematic on two fronts. First, deviating dramatically from the established and validated cost-effectiveness framework introduces risk that the PM and other decision-makers may not be willing to accept. Second, the concept of Pareto efficiency is well regarded as a means to reduce to complexity of the alternative space to a trade space that can be reasonably managed by decision-makers. Not only is this

المنارة للاستشارات
www.manaraa.com

reduction in complexity forfeited by the current representation, but substantial complexity is added in the form of alternative assumptions regarding the timing/selection of technology infusions. Consequently, some mechanism is required to the reduce the complexity of this formulation to a representation that would enable acquisitions professionals to make informed decisions on the selection of a TRP.

These observations allow Research Question Two to be refined as follows:

RQ2.1 What is an appropriate method to anticipate the evolving properties of maturing technology?

RQ2.2 What measures of performance would allow competing Technology Refresh Plans to be compared in such a way that the set of all feasible alternatives can be reduced to a more concise trade space?

RQ2.3 What is an efficient, automated procedure to find the efficient trade space of competing Technology Refresh Plans for an arbitrary design problem?

5.2 Technology Forecasting

Methods for predicting the evolving properties of maturing technology come from the greater domain of Technology Forecasting. The elements within this domain that are relevant to the problem at hand are those devoted to the prediction of how technology attributes improve over time. This section will review these methods in order to determine how intelligent predictions can be made during the technology refresh planning phase of OSA development.

5.2.1 Univariate Growth Models

Univariate forecasting methods are predicated on a class of functions collectively referred to as "S-Curves". S-Curves model the technological "growth" of technology based on an analogy to the stages of biological growth demonstrated in Figure 38 [45].

المنارة للاستشارات
www.manaraa.com

The embryonic stages reflects the period of time in which little is known about the underlying physics of the technology. Significant effort, often in the form of fundamental research, is therefore required to create modest improvements. Improvements become easier to achieve as knowledge accumulates over time, and eventually reach a point of critical mass where there is a dramatic increase in the rate of improvement. This development threshold marks the technology's entry into the growth phase. The rapid pace of development in the growth phase eventually reaches a point of decreasing marginal returns, at which point the technology is deemed to be "mature". Maturation continues until the technology approaches the limitations of the engineer's ability to exploit the underlying physics of the technology's operation. This marks the aging phase of technological development, where, similar to the embryonic stage, substantial time and effort are required to generate minor improvements.

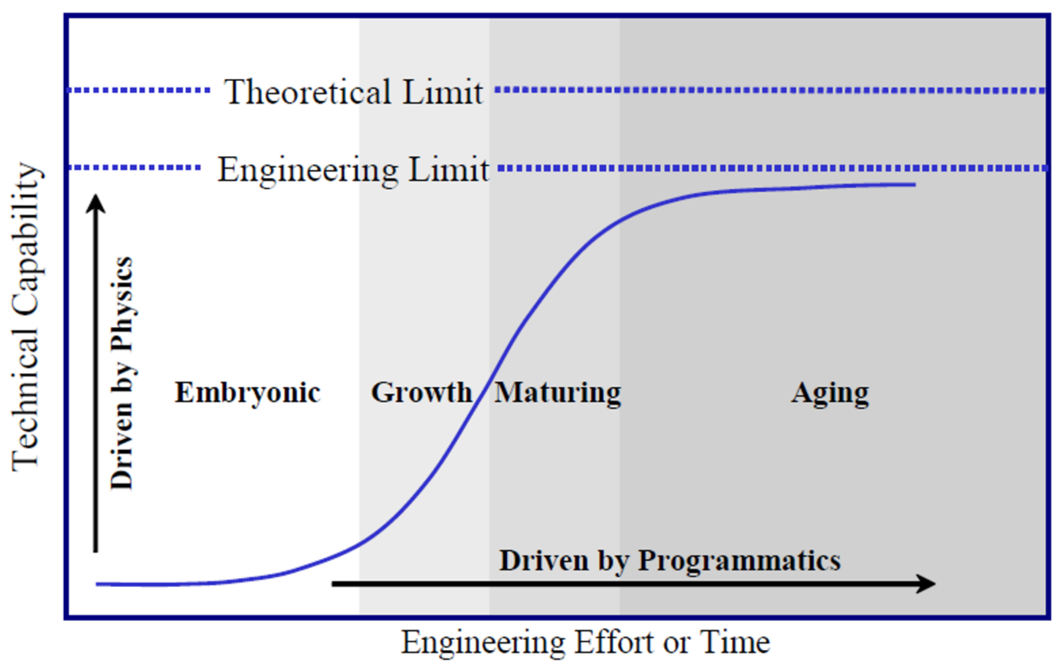


Figure 38: Analogy of Technology Growth as a Biological Process [45]

www.manaraa.com

Though univariate S-Curves may seem to have a rather dubious, heuristic heritage, they have proven to be quite accurate representations of real world technological advances. This approach has therefore been the subject of considerable academic research over the years, which led to the development of numerous functional forms to model the underlying evolutionary behavior. These methods can generally be grouped into one of two categories - absolute and relative. Absolute models quantify component performance, x_t, as a direct function of the independent parameter time, t. Relative models, on the hand, quantify the rate of change in a component's performance, dx_t, as a function of the most recently achieved performance level, x_{t-1}. Table 7 provides a summary of the most common functional forms found in the literature, where the parameter, L, represents the engineering limit of the underlying technology [45, 178].

Table 7: Common Univariate Growth Curves

Model Name	Equation	Type
Logistics [31]	$x_t = \frac{L}{1+\alpha e^{-\beta t}}$	Abs
Gompertz [111]	$x_t = Le^{-\beta t - \alpha t}$	Abs
Mansfield-Blackman [112, 22, 21]	$ln\left(\frac{x_t}{L-x_t}\right) = \beta_0 + \beta_1 t$	Abs
Linear Gompertz [177]	$ln\left(-ln\left[\frac{x_t}{L}\right]\right) = \beta_0 + \beta_1 t$	Abs
Weibull [151]	$ln\left(ln\left[\frac{L}{L-x_t}\right]\right) = \beta_0 + \beta_1 ln(t)$	Abs
Bass [18, 78, 168]	$x_t = \beta_0 + \beta_1 x_{t-1} + \beta_2 (x_{t-1})^2$	Rel
NSRL [63, 62]	$ln(x_t) = \beta_0 + \beta_1 ln(x_{t-1}) + \beta_2 ln(L - x_t - 1)$	Rel
Harvey [76]	$ln(x_t) = \beta_0 + \beta_1 t + \beta_2 ln(x_t - 1)$	Rel
Extended Riccati [102]	$\frac{x_t}{x_{t-1}} = \beta_0 \beta_1 x_{t-1} + \beta_2 \left(\frac{1}{x_{t-1}}\right) + \beta_3 ln(x_t - 1)$	Rel

Application of any of the methods listed in Table 7 requires the designer to regress

www.manaraa.com

the model against a set of time series performance data in order to identify the unknown parameters. Once those parameters are identified, the model can be extrapolated into the future to ascertain an estimate for the future evolution of the technology under consideration. Multiple forecasters have described a similar, albeit more comprehensive, approach to this process [45, 141, 76], which Twiss succinctly outlines as follows [172]:

1. Identify the appropriate attribute for the product of the system in which it is embedded
2. Determine the technology parameter by which the attribute can be measured
3. Collect data for the past progress of this parameter over time
4. Establish the natural/physical limit for the parameter using the technology being forecasted
5. Fit an S-Curve to the data which becomes asymptotic at the limiting level
6. Consider events or other trends which may affect the future development of the technology and thereby influence the shape of the curve, i.e. the emergence of a new technology or other factor which might affect the funding necessary to drive the advance

Unfortunately, the process outlined by Twiss is unlikely to ever provide a perfect prediction of the future. Extrapolative methods are, by their very nature, subject to uncertainty, as there is no way to predict the future with absolute certainty. Martino notes that there are three drivers of uncertainty in this process: historical data provides an inaccurate representation of the true state of the art, the growth curve is an imperfect model of the underlying process, and the prediction of the upper limit is incorrect [121].

المنارة للاستشارات
www.manaraa.com

There is no apparent mechanism with which to model the first driver of uncertainty. This work will therefore assume that imperfections in the available data are a source of risk in the process that one must assume in order to take advantage of the benefits of an OSA design. The second driver is an expression of modeling error, which modern statistical analysis software is able to integrate into analyses with ease. The more problematic driver lies with the estimation of the engineering limit. The upper limit is commonly estimated as a regression parameter alongside the regression coefficients [50, 178], which implies that error modeling software should be sufficient to determine a measure of confidence in the estimate. Martino warns against accepting this assumption, however, under the argument that the productivity of early technology development is only minimally influenced by the upper limit [121]. To illustrate this point, Martino conducted an experiment in which he varied the upper limit of steam engine efficiency from 45 to 55 percent. Using a Logistic regression, he found that the inflection point, i.e. the transition threshold for technological maturity, shifted from 1900 to 1925. Martino concluded that, "even a small error in the upper limit can result in a fairly significant error in the forecast" [121]. Thus, error in the limit estimation should be an area of significant consideration when formulating methods to evaluate uncertainty / flexibility in the next chapter.

A second, and potentially more significant, complication with univariate methods is that they can only consider one DV at a time. If a component is only governed by a single DV, then this point is moot. Yet, Martino notes that this scenario is somewhat rare [119]:

> In some cases the performance parameters are dominated by a single parameter sufficient to characterize the state of the art. In such cases the remaining technical parameters are essentially irrelevant from the standpoint of measuring the level of technology....in most cases, however, a

المنارة للاستشارات
www.manaraa.com

technological device can be characterized by a set of technological parameters which measure several characteristics.

It can be shown that if a component is governed by a generic vector of attributes, then univariate methods are appropriate if, and only if, the technological progression of attributes are mutually independent [45]. In other words, if allocating time and resources to improve one attribute has an impact on the progression of another (e.g. development under a limited RDT&E budget), then this assumption is invalid, and univariate methods are inappropriate. As this methodology is intended to be platform/component agnostic, it follows that it should be able to accommodate multiple, dependent DV. Thus, while univariate methods provide a foundation for more sophisticated approaches, they alone are insufficient for this methodology.

5.2.2 Scoring Models

The key capability that is lacking in univariate methods is the ability to simultaneously consider multiple attributes. Though there are volumes of research on the subtleties of univariate methods, there are only a handful of methods capable of considering multiple variables. The first attempt to resolve this gap started in the 1990's with the scoring model approach.

Scoring models can be described as somewhat of a hybrid approach, drawing from both the univariate and multivariate domains. The basic premise is that multiple attributes can be aggregated into a single technology measure, which is intended to be representative of the SOTA at that time. As time progresses and technology improves, the score of successive generations improves as well. Univariate methods can then be fit to the time history of these scores in order to extrapolate into the future [120]. Consequently, scoring models provide a means for the designer to propose a vector of desirable attributes, convert these parameters into a score, and then extrapolate to determine when such a technology could be available.

المنارة للاستشارات
www.manaraa.com

Martino provides a systematic procedure to develop a scoring model by decomposing variables of interest into three categories: overriding, traceable, and optional. Overriding variables are those whose absence would render the system worthless, and are therefore treated as a binary variable multiplying the remainder of the scoring function. If the variable is present, the score is unchanged; if the variable is absent, then the overall score is zero. Traceable variables simply represent those that would be identified in a Pareto analysis. Martino advocates that these should be normalized, and linearly combined under a subjective weighting scheme. Finally, optional variables are represented as a quantity $(1 + x)$, where x represents the added "value" provided by the variable at a given setting. This quantity is also a multiplier of the overall score, which provides the necessary property that the score is unmodified if the variable is absent [120].

By convention, scoring models are constructed as a ratio, where variables intended to be maximized are placed in the numerator and variables intended to be minimized are in the denominator. Equation 26 provides an example of a scoring model used as a predictor for fighter aircraft. Note that in this example, each of the terms in the numerator represents a group of independent tradeable variables, each of which has a distinct weighting scheme [121, 120].

$$Score = \frac{Maneuver \cdot Availability \cdot Range \cdot Payload \cdot Speed \cdot Avionics \cdot Weapons}{1 + Takeoff\ Roll} \tag{26}$$

While the scoring model is a conceptually simple way to manage the complexity of multiple variables, there are significant drawbacks. Perhaps the most obvious is the fact this approach is inherently subjective, as there is no consistent method to establish which variables should belong to each category, how they should be scaled, and what the appropriate weightings should be. This lack of consistency and traceability

المنارة للاستشارات
www.manaraa.com

was noted as one of the barriers to OSA implementation, making this approach unlikely to be acceptable in the eyes of the stakeholder. Even more significant, however, is that by collapsing all of the attributes into a single measure of performance, the scoring model loses any information regarding the interdependence between attributes. Without this information, it is impossible to recreate a Pareto surface defining the SOTA at some future time. This is the essential property that must be present in an acceptable forecasting method for this methodology. Consequently, scoring models represented an important milestone in the development of multivariate forecasting methods, but they are insufficient for the methodology under development.

5.2.3 Technology Frontiers

Scoring models were found to be lacking due to their requirement that multiple dimensions of technology attributes be collapsed into an aggregate SOTA measure, which forfeits the ability to evaluate the relationship between dimensions. Technology frontiers, on the other hand, preserve this multidimensional relationship in the forecasting process by modeling technological evolution in terms of the progression of the Pareto surface itself. The original work in this field is attributable to to Knight [98], who fit two dimensional Pareto surfaces to digital computers. This fit was done after a logarithmic transformation of the data, with one attribute serving as the dependent/independent variable. The result was a series of parallel lines that shifted across the logarithmic domain as time progressed. This work provided the motivation for the planar frontier approach proposed by Alexander and Nelson given by Equation 27, whereby a hyperplane is fit to the SOTA curves at different points in time [3]. In practical application, the intent is to simply substitute the vector of desired attributes into the equation to determine when the SOTA curve will intersect the corresponding point in the objective space.

المنارة للاستشارات
www.manaraa.com

$$t = \beta_1 x_1 + \beta_2 x_2 + \ldots \beta_n x_n \tag{27}$$

The challenge with the planar frontier approach lies in the assumption of a linear frontier. This assumption may prove valid in the logarithmic domain, but a more general approach would afford some degree of curvature to the trade space. Recognizing this, Dodson proposed the approach provided in Equation 28, in which the trade space is modeled with a ellipsoid rather than a plane[8] [58].

$$t = \sum_{i=1}^{n} \left(\frac{x_i}{c_i} \right)^2 \tag{28}$$

Here, n is the number of attributes; t is the introduction date for the corresponding component; x_i represents the i^{th} technology characteristic value; and c_i are the unknown intercepts determined through the regression [39]. The result of this approach at a given time is provided as the interior curve of Figure 39. Martino later modified this work to allow any even exponent in Equation 28 [119]. The result of raising this exponent is to increase the "squareness" of the resulting curve, which is depicted by the remaining curves in Figure 39.

5.2.4 Data Envelopment Analysis

Technology Forecasting through Data Envelopment Analysis (TFDEA) is an extension of a business planning technique known as Data Envelopment Analysis (DEA). The model was originally developed for industrial plants, where varying levels of inputs can be used to produce different levels of output. The purpose of this method is to determine the optimal settings such that the greatest quantity of a collection of outputs, which are known as Decision Making Units (DMU) in the DEA vernacular, can be generated with the least amount of inputs. The appropriate settings for

[8]Equation 28 is a slight modification of the original formulation provided by Cole, which yields a more direct representation of an evolving trade space than the radial measurement process originally conceived by Dodson [39].

المنارة للاستشارات
www.manaraa.com

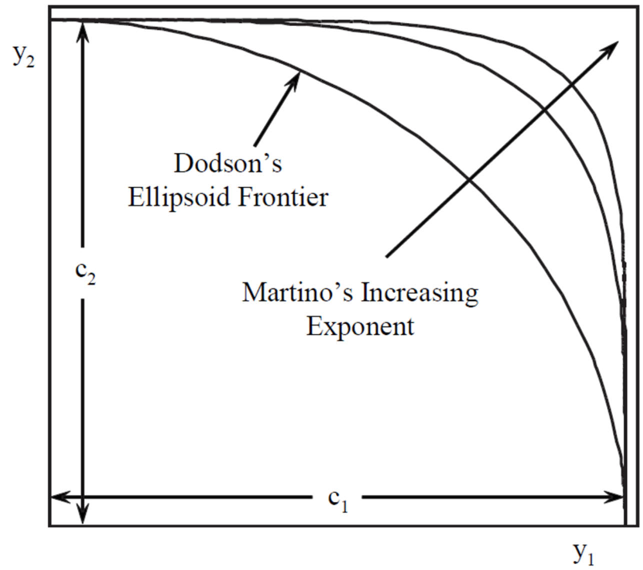


Figure 39: Notional Ellipsoid Frontiers [45]

optimal efficiency are found through the solution of the linear programming problem provided in Equation 29 [39].

$$
\begin{array}{lll}
max & \theta_k = \frac{\sum\limits_i \mu_i y_{r,k}}{\sum\limits_i \nu_i y_{r,k}} & \\
subject\ to & \frac{\sum\limits_{r=1}^{s} \mu_r y_{r,j}}{\sum\limits_{i=1}^{m} \nu_i x_{i,j}} & \forall\ j\ \in \{1, 2, \ldots, n\} \\
& \mu_r, \nu_i\ >\ 0 &
\end{array}
\tag{29}
$$

Equation 29 is interpreted as follows: θ_k represents the efficiency of the k^{th} DMU[9]; n is the number of DMU; s is the number of outputs; m is the number of inputs; x and y are the respective inputs and outputs. Figure 40 provides a flowchart of the adaptation of this approach into the TFDEA methodology.

The outer loop of Figure 40 is a loop iterating across a discrete set of time periods of interest, t_0 to T, where t_f is the period to which technology data from all other periods is projected. The inner loop iterates over each DMU (i.e. technology data point) in the database between t_0 and the current value of t_f. The interior portion of

[9]Note that 1.0 is the greatest possible efficiency.

المنارة للاستشارات
www.manaraa.com

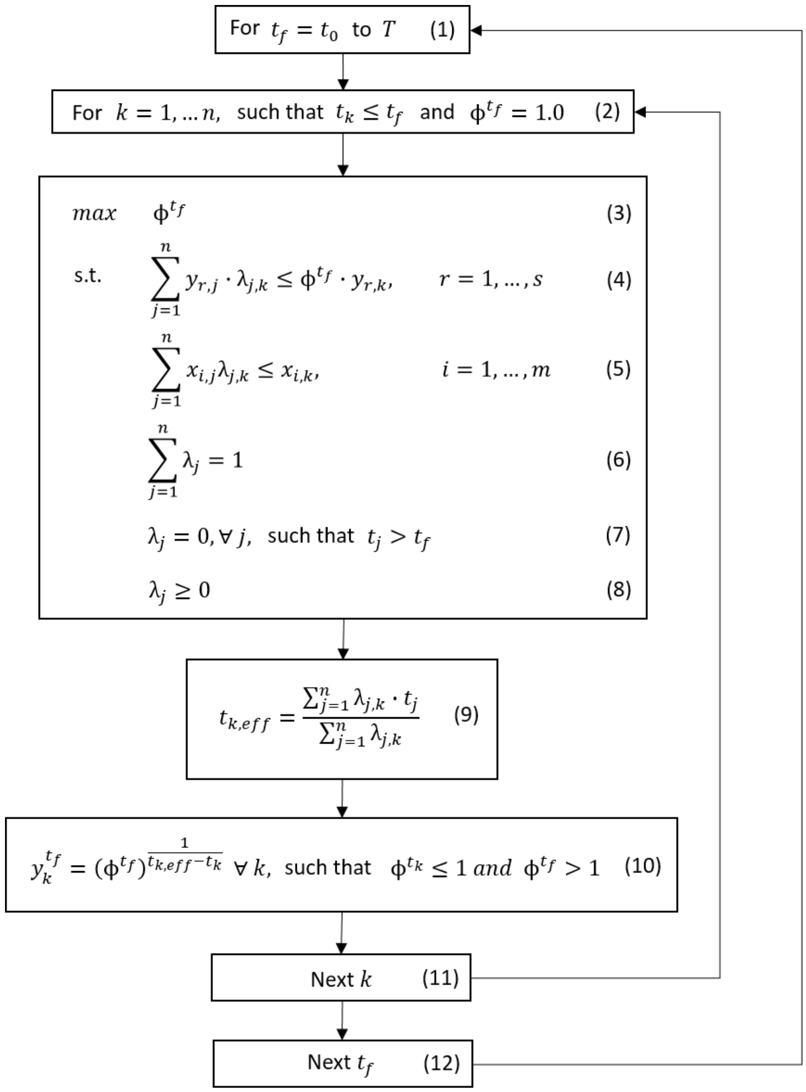


Figure 40: TFDEA Algorithm [89]

the algorithm then solves a linear programming problem to determine the efficient, piece-wise linear frontier at each point in time. The rate of change between these linear frontiers is assumed to be constant. Therefore, when a new vector of attributes is proposed, TFDEA simply calculates the time it will take, at a constant rate, for the frontier to expand that to point in the design space [39]. Figure 41 provides a simpler visual interpretation of this process [105], and the interested reader is referred to Inman's original work for a more in-depth description [89].

Though somewhat complicated in its formulation, TFDEA has proven to be widely

www.manaraa.com

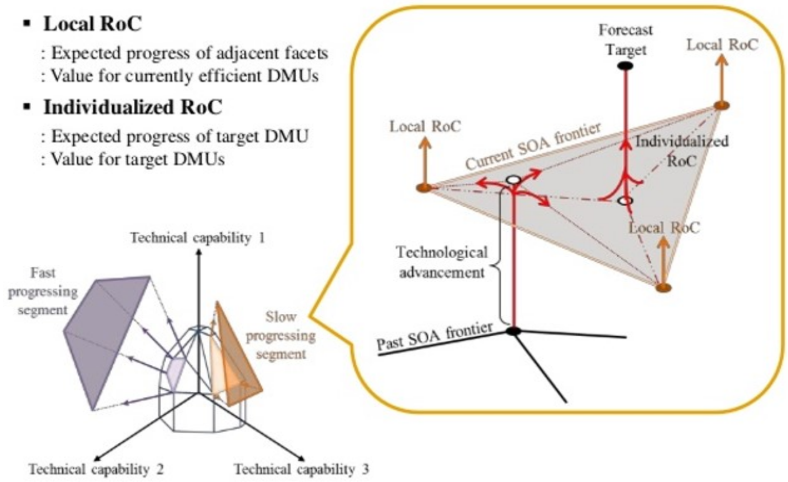


Figure 41: Visualization of TFDEA Application [105]

popular in academic literature. Various authors have successfully applied this approach to database systems, microprocessors, hard disk drives, and portable flash storage [6, 7]. In addition, Inman directly compared his method to the technology frontier approach in predicting the entry date of military aircraft, and determined that TFDEA has better performance in terms of both mean square and absolute deviations [90]. This approach is therefore worthy of further consideration.

5.2.5 Multi-Dimensional Growth Models

The final mechanism for multi-variate forecasting is the Multi-Dimensional Growth Model (MDGM) proposed by Danner [45]. The basic premise of the formulation starts with the initial understanding of technology frontiers and univariate S-Curves previously established. Recall that each curve within the technology frontier plot represents feasible combinations of technical capabilities that can be achieved at any single point in time. This observation is shown graphically in Figure 42 for an instance in which a technology is defined by two attributes, both of which are intended to be maximized.

MDGM then imposes an additional assumption on this relationship [45]:

www.manaraa.com

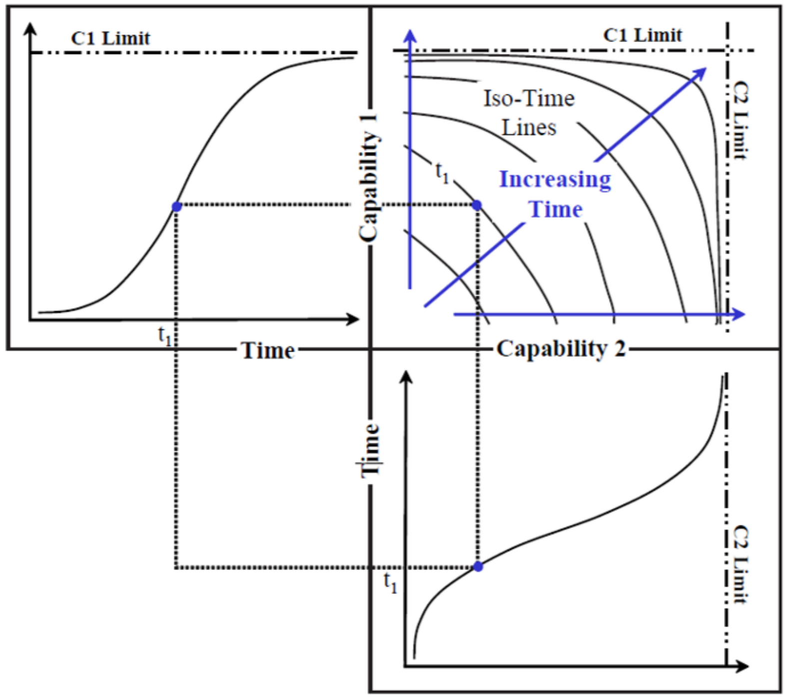


Figure 42: Technology Frontier and S-Curve Relationship [45]

> The feasible levels of capability that can be achieved by any one attribute of a complex technology advance over time according to a technology S-curve provided all other attributes remain constant.

There is an intuitive appeal in this assumption. For example, if an entity wishes to move beyond the state of the art in a technology with two attributes, then some measure of effort must be expended. If all effort is directed toward one attribute, then it is reasonable to expect that one attribute will proceed along its S-Curve as predicted, while the second attribute will remain constant. This concept is depicted in Figure 43.

If, on the other hand, the entity's effort is split between improving both attributes, then that same measure of effort must be divided among the competing attributes. This implies that the first attribute will not progress as far along its corresponding univariate S-Curve as it would under the previous assumption. This concept is depicted in Figure 44.

المنارة للاستشارات
www.manaraa.com

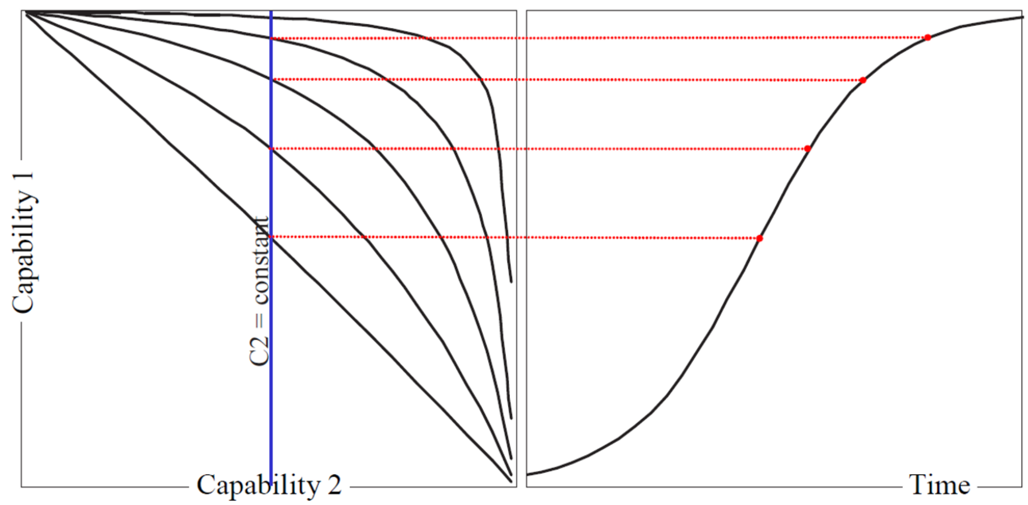


Figure 43: Constant Parameter Relationship [45]

The MDGM methodology also assumes that the limit of each attribute is not influenced by the level of capability of the remaining attributes. Under this assumption, the variability imposed on the univariate S-Curve of a given attribute can be modeled as a left or right shift, where the magnitude of that shift is an unknown function of the remaining attributes. Assuming the S-Curves are modeled through the standard Logistics function, this assumption can be analytically imposed through Equation 30 below:

$$\begin{aligned} x_1 &= \frac{L_1}{1+a_1 e^{-b_1[t-f(x_2)]}} \\ x_2 &= \frac{L_2}{1+a_2 e^{-b_2[t-f(x_1)]}} \end{aligned} \tag{30}$$

Finally, it is further assumed that each of the i univariate curves start at $x_{i,0}$, as opposed to zero. With these assumptions in place, the unknown functions can be found by solving both relationships given in Equation 30 for time, t, and setting the resulting expressions equal to one another. The terms on the opposing side of the equality must then correspond to the unknown functions. If one substitutes this result back into the formula for time and simplifies, then the result is given by Equation 31. Finally, it can also be shown that the result in Equation 31 can be extended to n

المنارة للاستشارات
www.manaraa.com

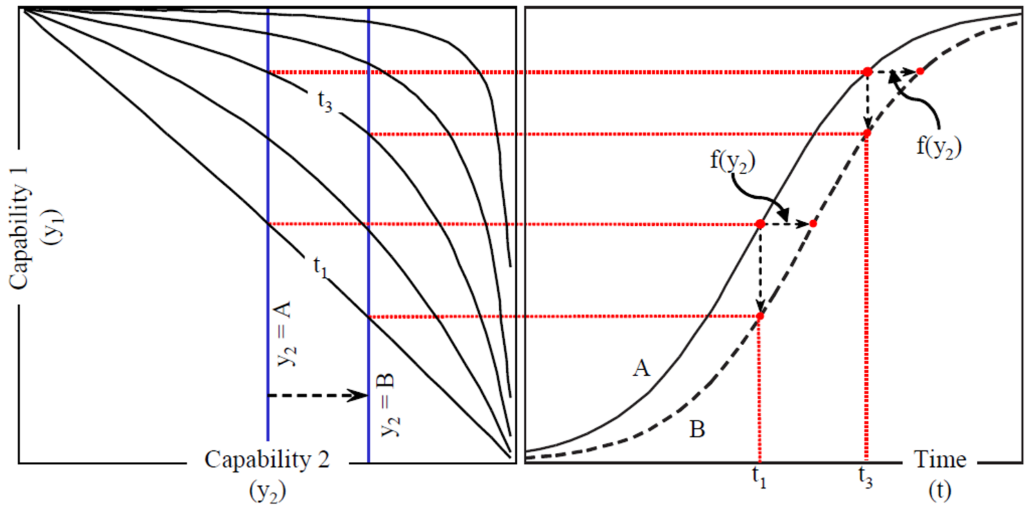


Figure 44: Impact of an Additional Attribute on an Attribute-Specific S-Curve [45]

dimensions using Equation 32 [45].

$$t = a + \beta_1 ln\left(\frac{L_1 - x_1}{x_1 - x_{1,0}}\right) + \beta_2 ln\left(\frac{L_2 - x_2}{x_2 - x_{2,0}}\right) \tag{31}$$

$$t = a + \sum_{i=1}^{n} \beta_i ln\left(\frac{L_i - x_i}{x_i - x_{0,i}}\right) \tag{32}$$

5.2.6 Downselection of Alternative Methods

Three of the methods previously described provide the necessary multivariate capabilities for the forecasting mechanism required in this methodology: Technology Frontiers [58], TFDEA [89], and MDGM [45]. Moreover, all three have the advantage of being objective, as the forecast results stem only from a database of component performance and the assumptions clearly articulated in their formulation. Selecting a method from within this family should therefore rest on their accuracy.

In his thesis, Cole performed such a study [39]. To accomplish this, he conducted a full factorial set of experiments in which each model was applied to evaluate the datasets used in the experimental validation of the other methods. Null hypotheses

المنارة للاستشارات
www.manaraa.com

were formulated for each data set that the other methods would have a lower absolute deviation, and a paired t-test was applied to each comparison in order to reject the hypothesis. The results of this analysis are provided in Table 8. In all three data sets, the hypothesis that the Technology Frontier had lower mean error than the other models was rejected with 95% confidence. The results were far less clear when comparing TFDEA and MDGM.

Table 8: Statistical Results of Attempted Predictions of Date of Introduction [39]

	t-value	p-value
Jet Engine Data - Last 6 Values Predicted		
$H_0 : \mu_{MDGM}^{abs.deviation} - \mu_{TFDEA}^{abs.deviation} > 0$	1.09	0.34
$H_0 : \mu_{MDGM}^{abs.deviation} - \mu_{Tech.Frontier}^{abs.deviation} > 0$	3.33	0.01
$H_0 : \mu_{MDGM}^{abs.deviation} = 0$	1.21	0.28
$H_0 : \mu_{TFDEA}^{abs.deviation} = 0$	0.37	0.73
Microchip Data - Last 14 Values Predicted		
$H_0 : \mu_{MDGM}^{abs.deviation} - \mu_{TFDEA}^{abs.deviation} > 0$	2.71	0.01
$H_0 : \mu_{MDGM}^{abs.deviation} - \mu_{Tech.Frontier}^{abs.deviation} > 0$	1.79	0.05
$H_0 : \mu_{MDGM}^{abs.deviation} = 0$	0.97	0.35
$H_0 : \mu_{TFDEA}^{abs.deviation} = 0$	1.11	0.29
Jet Fighter Data - Last 7 Values Predicted		
$H_0 : \mu_{MDGM}^{abs.deviation} - \mu_{TFDEA}^{abs.deviation} > 0$	0.35	0.37
$H_0 : \mu_{MDGM}^{abs.deviation} - \mu_{Tech.Frontier}^{abs.deviation} > 0$	6.07	0.00
$H_0 : \mu_{MDGM}^{abs.deviation} = 0$	1.60	0.16
$H_0 : \mu_{TFDEA}^{abs.deviation} = 0$	0.33	0.75

Cole also evaluated the residuals of each experiment in terms of their similarity to a Normal distribution. He concluded that the Technology Frontier approach not only lags behind TFDEA and MDGM in accuracy, but also appeared to have non-normally distributed errors. TFDEA and MDGM, however, do not have differences that are statistically significant (both had a mean error of approximately two years with a time unit of integral years), and exhibited normally distributed errors. Both

المنارة للاستشارات
www.manaraa.com

models therefore appeared to have equivalent performance, and were deemed to be equally likely to perform well on a given problem [39].

With these results in mind, this work will proceed with the MDGM as the forecasting platform. The reason for this is that, unlike TFDEA, MDGM provides a *closed form* expression for the SOTA at a given point in the future. As the next section will demonstrate, this is a powerful property that will facilitate the automated exploration of a large trade space for TRP's.

5.2.7 Novel Approach to Generate Technology Refresh Plans

Each of the multivariate methods considered in the previous section is intended to be used to address questions of the following form: *given a vector of desirable attributes, when would such a component be available?* The problem derived in the development of technology refresh planning requires the inverse question to be addressed: *Given a future date, what combinations of component attributes will be feasible?* It is unclear how TFDEA could be adapted to answer the inverse question, but a simple transformation can be applied to the MDGM that allows it to resolve this gap. Consider again the relationship between time and a generic design vector expressed by Equation 32. Solving this equation for any one of the n DV variables results in Equation 33, which expresses the value that DV would take given the time period under consideration, t, and the settings of the $n-1$ remaining DV. Also, note that the selection of which of the DV serves as the dependent variable is arbitrary; the n^{th} variable was simply chosen for the sake of mathematical brevity.

$$x_n = \frac{L_n - x_{0,n}}{1 + e^{-b_n\left[t - a + \sum_{i=1}^{n-1} \frac{1}{b_i} ln\left(\frac{L_i - x_i}{x_i - x_{0,i}}\right)\right]}} + x_{0,n} \tag{33}$$

Figure 45 illustrates the significance of this result by applying the formulation provided in Equation 33 to the original data-set used to demonstrate the MDGM approach. For the sake of simplicity, the author normalized this data under the

www.manaraa.com

convention provided in Table 9. Note that this convention dictates that the utopia point occurs where x and y are maximized, and z is minimized. This point would be found in the lower right hand corner of each plot. As one would expect, the frontier starts as a concave hull with respect to the utopia point. Over time, however, this shape gradually inverts into a concave shape driving ever closer to the utopia point.

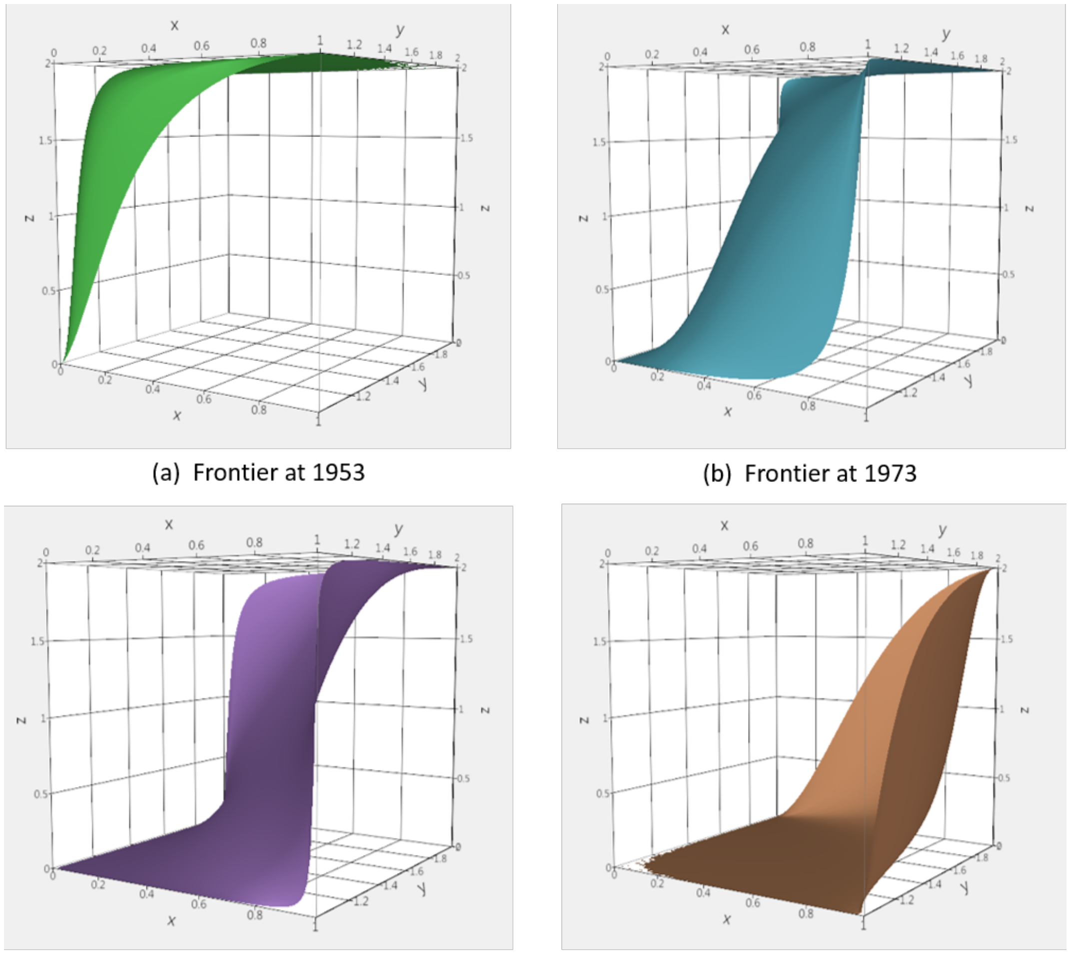
Table 9: Normalized Formulation of Original MDGM Data [45]

Variable	Start	Limit	Quality Characteristic
x	0	1	Maximize
y	1	2	Maximize
z	2	0	Minimize

With this in mind, reflect back on the depiction provided by Figure 35. Assuming that an MDGM was fit to the relevant data, which includes both the limits and starting values, then the frontier at a future point in time can be determined by substituting the desired time, t, into Equation 33. New design vectors are then established based on whatever point is selected on the corresponding frontiers, and the effects of this change propagate through M&S as required by Equation 25 to create a change in RV. The change in RV over time allows the designer to formulate the performance profile defining the TRP.

The only caveat in this process is that technology infusions do not happen overnight. It will require some amount of development time to decide that an upgrade is warranted, determine which component should be infused, and then conduct the requisite RDT&E / logistical deployment of the upgrade. Though OSA designs are developed in such a way that this process is significantly shorter than that of an equivalent closed architecture, it is still likely that this delay will be non-negligible. Assuming that the field of commercial components under consideration is frozen at the start of this process, the existence of a delay implies that the effective date of the technology inserted into the OSA platform will be given by Equation 34. In this formulation,

المنارة للاستشارات
www.manaraa.com



(a) Frontier at 1953 (b) Frontier at 1973

(c) Frontier at 1984 (d) Frontier at 1999

Figure 45: Pareto Visualization of MDGM Results at Successive Time Intervals

$t_{upgrade}$ is the date in which the upgrade is available to the end user and t_{dev} is the length of time required to executed the development process, which is assumed to be a known parameter. Substituting this relationship into the original closed form solution for the SOTA curve at time t then yields Equation 35. This is the final form of the MDGM that will be leveraged in this work.

$$t_{eff} = t_{upgrade} - t_{dev} \tag{34}$$

$$x_n = \frac{L_n - x_{0,n}}{1 + e^{-b_n\left[t_{eff} - a + \sum\limits_{i=1}^{n-1} \frac{1}{b_i} ln\left(\frac{L_i - x_i}{x_i - x_{0,i}}\right)\right]}} + x_{0,n} \tag{35}$$

المنارة للاستشارات
www.manaraa.com

In review, the derived requirements from the previous discussion are satisfied by the modified MDGM approach, and Research Question 2.1 is therefore assumed to be satisfied by this method. Yet, it remains unclear how one should determine the timing of upgrades, and, assuming this timing is known, what design points should be selected on the corresponding frontiers. Addressing these questions will, however, require a method to compare and evaluate TRP's. This capability is lacking at the moment, but the next sections will generate a mechanism to facilitate this comparison in the course of addressing Research Questions 2.2 and 2.3.

Finally, the previous discussion clearly indicates that the MDGM is an appropriate means to satisfy the needs of the third step in the proposed methodology. Figure 46 updates the process model with this observation.

5.3 Comparison of Competing Technology Refresh Plans

It was previously established that cost-effectiveness analyses for closed architectures require the decision-makers to determine whether an incremental increase in effectiveness was justified by the corresponding cost. According to the official Air Force AoA Handbook, "there is no formula for doing this; it is an art whose practice benefits from experience" [135]. A challenge with the methodology up to this point is that the inclusion of time as an additional degree of freedom exponentially increases the complexity of the analysis. To demonstrate this point, consider a thought experiment in which the period of time under consideration covers 20 years, and that it takes approximately 2 years to perform the required RDT&E and logistic deployment of an upgrade. This scenario implies that there are 10 potential upgrade opportunities, which equates to 1024 (i.e. 2^{10}) possible combinations of upgrades. If, on the other hand, the program could theoretically accommodate an annual refresh cycle, then the number of possible upgrade opportunities swells to 1,048,576 (i.e. 2^{20}). No matter how experienced an acquisitions decision-maker is in the practice of their art,

www.manaraa.com

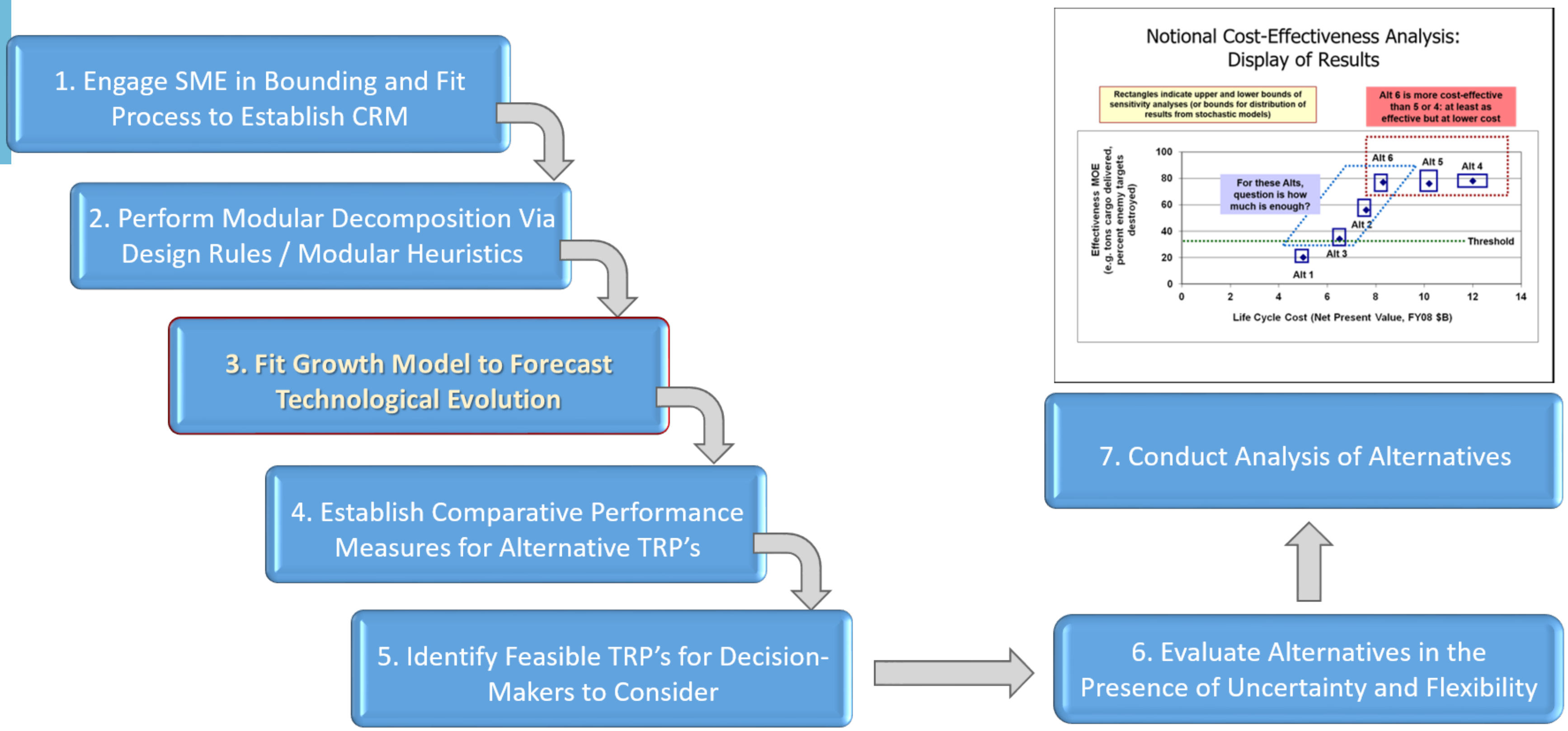


Figure 46: Methodology Update: Step Three - Growth Model as the Forecast Mechanism for Technological Evolution

المنارة للاستشارات
www.manaraa.com

this is clearly a degree of complexity that cannot be reasonably managed. Yet, the methodology at this point would require this level of analysis to properly consider the entirety of the trade space.

Section 5.1 noted that Pareto optimality was the approach used to reduce the complexity of the design space for traditional system configurations. The logical question then, is what does Pareto optimality mean in the context of an OSA? One obvious example could be constructed in which one effectiveness profile is entirely contained within another, such as the example provided on the left side of Figure 47. If the greater profile were available at a lower cost, then it could be considered efficient compared to the lower frontier because it provides greater effectiveness at all times, and at a lower cost. It is not clear, however, how this principle could be applied to the scenario depicted on the right side of Figure 47, where one plan provides a larger overall increase in performance, while the other provides the improvement earlier in the life cycle.

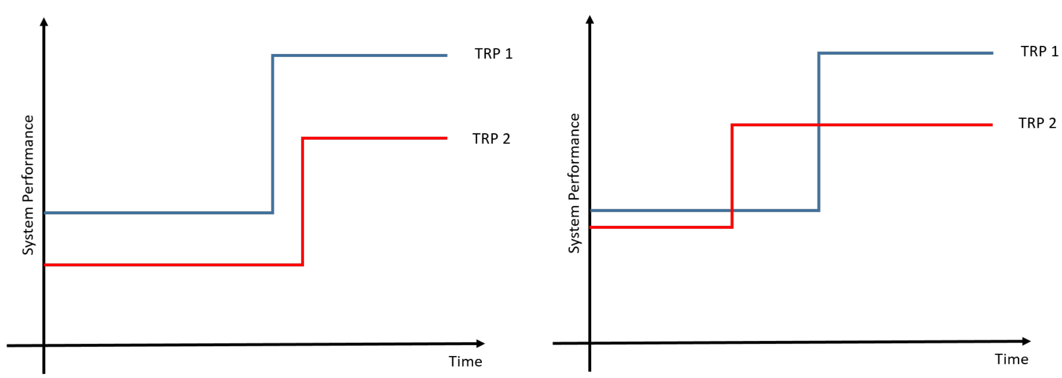


Figure 47: Application of the Pareto Analogy to Effectiveness Profiles

The conclusion of this thought experiment is that there is simply too much complexity in the formulation of TRP's for a purely *a posteriori* articulation of preferences. Some additional mechanism is required to reduce the trade space to a reasonable extent.With this in mind, the question to be resolved is what would lead a decision-maker to prefer one performance profile over another. The answer is clearly

www.manaraa.com

not self-evident, but a useful starting point in attempting to answer this question is to consider what would lead one to select a given closed architecture system over another. Unlike the OSA context, this is an area which has received considerable academic attention in the form of Value Theory.

5.3.1 Value Theory

Value theory can be informally defined as a sub-domain in the broader field of utility theory where outcomes are entirely deterministic (i.e. uncertainty is negligible) [110, 47, 95]. At its core, value theory, and by extension utility theory, is a normative modeling tool. This means that the *a priori* articulation of preferences is not meant to mimic the choices of human decision-makers, but to help humans make better decisions [167]. The need for such methods stems from the observation that unaided decision making, particularly in the context of complex problems, often exhibits inconsistencies, irrationality, and sub-optimal choices [171]. To address these challenges, value theory defines a set of "axioms of rational behavior" enumerated below[10] [175]. These axioms use the following notation: $\mathbf{y_1}, ..., \mathbf{y_n}$ are the different alternatives under consideration, $\mathbf{y_1} \succ \mathbf{y_2}$ states that $\mathbf{y_1}$ "is preferred to" $\mathbf{y_2}$, and $\mathbf{y_1} \asymp \mathbf{y_2}$ states that $\mathbf{y_1}$ "is indifferent to" $\mathbf{y_2}$.

Definition (Completeness Axiom) *Either* $\mathbf{y_1} \succ \mathbf{y_2}$ *or* $\mathbf{y_2} \succ \mathbf{y_1}$ *or* $\mathbf{y_1} \asymp \mathbf{y_2}$

Definition (Transitivity Axiom) *If* $\mathbf{y_1} \succ \mathbf{y_2}$ *and* $\mathbf{y_2} \succ \mathbf{y_3}$ *then* $\mathbf{y_1} \succ \mathbf{y_3}$

Definition (Monotonicity Axiom) *The decision-maker's preferences over the range of an attribute are either monotonically increasing or decreasing*

With this in mind, a function $v : \Re^k \rightarrow \Re$ that satisfies the aforementioned axioms such that $\mathbf{y_1} \succ \mathbf{y_2} \succ .... \succ \mathbf{y_n} \Longleftrightarrow v(\mathbf{y_1}) > v(\mathbf{y_2}) > ... > v(\mathbf{y_n})$ is referred to as

[10] von Neumann and Morgenstern actually postulate 6 axioms, but the latter 3 pertain to uncertainty. As the present subject is value theory, those axioms are omitted [175].

المنارة للاستشارات
www.manaraa.com

a *value function*[11]. An important feature of value functions is that the numerical results of $\upsilon(\mathbf{y})$ are on an interval scale, which means the function only provides an ordering of outcomes. For example, if $\upsilon(\mathbf{y_1}) = 2$ and $\upsilon(\mathbf{y_2}) = 1$, then $\mathbf{y_1} \succ \mathbf{y_2}$, but the same relation holds when $\upsilon(\mathbf{y_1}) = 5$ and $\upsilon(\mathbf{y_2}) = 1$. This demonstrates that value functions can determine whether an alternative is preferable to another, but they cannot quantify how much better the superior alternative is to the inferior alternative. Moreover, it can be shown that the preference structure provided by a given value function is indifferent to a positive, affine transformation of the form $\widehat{\upsilon} = \alpha\upsilon + \beta$, where α and β are constants, and α is positive. This property is known as the *scale indifference* of preferences [77].

Another observation of the axioms of rational behavior is that they do not provide guidance as to how value functions should be constructed, nor is there any evidence to suggest that a single functional form would be applicable to all problems. Creating a value function therefore requires the designer to formulate a set of assumptions regarding the underlying preference structure before any particular functional form can be determined. By far the most common assumption in this domain is *preferential independence*. Broadly speaking, preferential independence states that the rank ordering of preferences for one RV does not depend on the values of the other RV's when the other RV's are held constant. In other words, if A and B represent particular values of y_1 and y_2, then y_1 is preferentially independent of y_2 when $\{y_{1A}, y_{2A}\} \succ \{y_{1B}, y_{2A}\}$ and $\{y_{1A}, y_{2B}\} \succ \{y_{1B}, y_{2B}\}$ [167].

There are two important results stemming from the assumption of preferential independence. First, if all pairs of RV are preferentially independent of their complementary subsets of RV (e.g. if y_1, y_2 is preferentially independent of $y_3, ..., y_k$), then all RV are mutually preferentially independent. Second, if all attributes are mutually

[11]The objective of design in this context is typically to maximize value. Therefore, the terms "value function" and "objective function" will be used interchangeably in this work.

المنارة للاستشارات
www.manaraa.com

preferentially independent, then there exists a value function of the form given by Equation 36. In addition, because a logarithmic transformation does not impact the resulting preference structure, the existence of Equation 36 guarantees the existence of Equation 37 [47, 95].

$$\upsilon\left(y_1, y_2, ..., y_k\right) = \sum_{i=1}^{k} \upsilon_i\left(y_i\right) \tag{36}$$

$$\upsilon\left(y_1, y_2, ..., y_k\right) = \prod_{i=1}^{k} \upsilon_i\left(y_i\right) \tag{37}$$

5.3.1.1 Value Functions

The previous section established that preferential independence among attributes guarantees the existence of both an additive and multiplicative form of an aggregate value function. The individual terms in the summation and product, respectively, are the dependent variable of a univariate scaling function, where the independent variable, the quantity being scaled, is the individual RV. However, this observation does not prescribe a specific functional form for either the aggregate value function, or the individual sub-functions. Table 10 lists some of the more common forms identified in surveys by various authors, but it is ultimately left to the designer to determine the particular form that best captures their preferences. Note that in this notation, G and B represent generic "good" and "bad" values of the corresponding RV, and the w_i represent the relative importance of a given RV toward the overall design objective.

5.3.1.2 Determining Relative Weights

One of practices apparent in the common value functions provided in Table 10 is the use of weight parameters, w_i, to scale individual sub-functions according to their relative significance. Determining these weights in a consistent, objective fashion,

المنارة للاستشارات
www.manaraa.com

Table 10: Common Aggregate Value Functions

Name	Value Function
Absolute Value Methods	$\min_{\mathbf{x}\in D} \quad V(\mathbf{x}) = \sum_{i=1}^{k} w_i \lvert y_i(\mathbf{x}) - G_i \rvert$
Weighted Square Sum	$\min_{\mathbf{x}\in D} \quad V(\mathbf{x}) = \sum_{i=1}^{k} w_i (y_i(\mathbf{x}) - G_i)^2$
Weighted Maximum	$\min_{\mathbf{x}\in D} \quad V(\mathbf{x}) = \max_{i} \frac{y_i(\mathbf{x}) - G_i}{B_i - G_i}$
Substitute Objective Function	$\min_{\mathbf{x}\in D} \quad V(\mathbf{x}) = \prod_{i=1}^{k} \frac{y_{i,max} - y_i(\mathbf{x})}{y_{i,max} - y_{i,min}}$
Kresselmeir-Steinhauser Function	$\min_{\mathbf{x}\in D} \quad V(\mathbf{x}) = \frac{1}{\rho}\ln \sum_{i=1}^{k} \; exp(\rho y_i(\mathbf{x}) - y_{i,max})$
Distance from Utopia Point	$\min_{\mathbf{x}\in D} \quad V(\mathbf{x}x) = \sum_{i=1}^{k} \; w_i (y_{i,max} - y_i(\mathbf{x}))^2$
Distance from Anti-Utopia Point	$\min_{\mathbf{x}\in D} \quad V(\mathbf{x}) = -\sum_{i=1}^{k} \; w_i (y_i(\mathbf{x}) - y_{i,min})^2$
Exponential Weighted Method	$\min_{\mathbf{x}\in D} \quad V(\mathbf{x}) = \sum_{i=1}^{k} \; w_i exp(c_i y_i(\mathbf{x}))$
Weighted Compromise Programming	$\min_{\mathbf{x}\in D} \quad V(\mathbf{x}) = \sum_{i=1}^{k} \; w_i (y_i(\mathbf{x}))^{c_i}$

www.manaraa.com

however, typically proves to be a difficult endeavor. Messac describes the general method for determining relative weights in practical applications as the set of iterative design loops shown in Figure 48 [123]. The first step in this iteration requires the decision-makers to express what they believe their preferences between RV to be. An analyst selects a set of weights reflecting these preferences, and obtains a potentially optimal design through M&S. If the design produced from this process fails to accurately reflect the decision-makers' preferences, new weights are chosen and the process repeats (inner loop). If the design passes this phase, the final question is whether the original preference statement led to an acceptable design. Failure means that the entire process must be redone (outer loop).

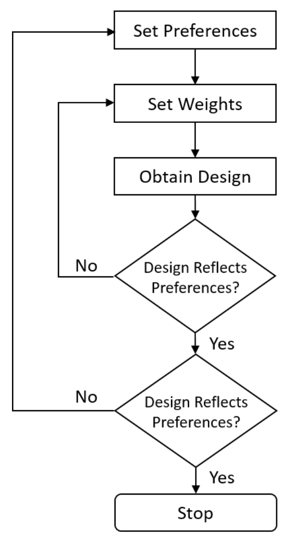


Figure 48: Weight Based Design Process

The first step in the design process depicted in Figure 48, setting preferences between RV, is perhaps the most challenging as it is inherently subjective. Saaty proposed the qualitative scale enumerated in Table 11 in an effort to standardize this process. This scale is used to evaluate the $\frac{1}{2}k(k-1)$ pairs of RV, which are recorded in an attribute matrix, A, with the reciprocal properties that $a_{ij} = \frac{1}{a_{ji}}$ and $a_{ii} = 1 \ \forall \ i$. For example, if y_1 is weakly more important that y_2, then $a_{12} = 3$ and $a_{21} = \frac{1}{3}$. With that in mind, there is no explicit need to use this scale. Designers

المنارة للاستشارات
www.manaraa.com

Table 11: Saaty's Pairwise Preference Scale [147]

Intensity	Definition	Explanation
1	Equal	Two criteria contribute equally to the design objective
3	Weak	Experience and judgment slightly favor one over the other
5	Strong	Experience and judgment strongly favor one over the other
7	Demonstrated	Criterion is strongly favored; dominance demonstrated in practice
9	Absolute	Evidence favoring one criterion is of the highest possible order

and decision-makers are free to use any scale deemed appropriate for the context of the problem under consideration, so long as the results are captured in the reciprocal attribute matrix [147].

Chu notes that there are two methods available to directly extract a set of weights from the attribute matrix: the Eigenvector Method and the Weighted Least Squares Method [38]. As the name implies, the Eigenvector method determines the eigenvalues, λ, associated with A by solving the familiar system of equations given by the kernel of $Det(A - \lambda I)$. The eigenvector corresponding to the largest eigenvalue is then assumed to be the vector of weights to be used in the aggregate value function. Weighted Least Squares is a less computationally intensive approach. The basic premise is that the entries of A, a_{ij}, should equate to w_i/w_j, which implies that w_i can be determined from the constrained optimization problem provided in Equation 38. In order to to minimize z, the Lagrangian is formed according to Equation 39. Differentiating the Lagrangian with respect to the weight parameters yields the set of $n+1$ linear, homogeneous equations with $n+1$ unknown weight parameters provided in Equation 40. The solution of this system provides the requisite information [85].

www.manaraa.com

$$
\begin{aligned}
&\text{minimize} \quad z = \sum_{i=1}^{k}\sum_{j=1}^{k}(a_{ij}w_j - w_i)^2 \\
&\text{subject to} \quad \sum_{i=1}^{k} w_i = 1
\end{aligned} \tag{38}
$$

$$
L = \sum_{i=1}^{k}\sum_{j=1}^{k}(a_{ij}w_j - w_i)^2 + 2\lambda\left(\sum_{i=1}^{k} w_i - 1\right) \tag{39}
$$

$$
\begin{aligned}
&\sum_{i=1}^{k}(a_{il}w_l - w_i)a_{il} - \sum_{j=1}^{k}(a_{ij}w_i - w_l) + \lambda = 0, \quad l = 1, \ldots, k \\
&\sum_{i=1}^{k} w_i = 1
\end{aligned} \tag{40}
$$

In summary, though there is no exact mechanism to consistently generate an "optimal" value function, the broader design community generally agrees that complex problems should be structured to ensure the assumption of preferential independence is satisfied. Satisfying this principle allows the value function to be decomposed into two elements. The first element is a set of scaling functions to transforms the raw value of a given RV into a performance measure of relative closeness to some design goal for that RV. The second is a set of weight factors governing the relative importance of a given RV to the overall design objective. These weight factors allow the individual RV scores to be combined into the overall score. Though conceptually simpler than the partial ordering approach, the shortcoming of the value function methodology is that it requires subjective information that may be inaccurate or unavailable in the early phases of design. Moreover, the group dynamics and conflicting preferences of decision-makers may make the proper formulation of a value function impossible, regardless of the technique chosen. Thus, care should be exercised when applying this approach in isolation.

www.manaraa.com

5.3.2 Dynamic Value

The methods available in classical Value Theory provide insight as to how one would aggregate multiple MoE into an overall value of aggregate effectiveness, but they do not address the fundamental question currently under consideration. The reason for this statement is that none of the methods provided in Table 10 consider the time dimension in value. For example, if the Substitute Objective Function were applied in an OSA context, then the functional form would be reflected in Equation 41. In this case, the MDGM would provide a mechanism to determine, given appropriate assumptions regarding the timing and selection of technology, the output of this function over time. Yet, it remains unclear how one would make intelligent decisions regarding which set of assumptions are preferable/efficient in the presence of variation over time.

$$V(\mathbf{x}, t) = \prod_{i=1}^{k} \frac{y_{i,max}(t) - y_i(\mathbf{x}, t)}{y_{i,max}(t) - y_{i,min}(t)} \tag{41}$$

A solution to this problem can potentially be found in a sub-domain of the Product Platform Design (PPD) literature known as Design for Adaptability (DFA). The DFA methodology was developed by Browning and Engel on a set of assumptions and principles closely aligned with those previously established for OSA design, albeit in a commercial context[12] [26]. In the development of their methodology, the authors extended the fundamental concepts of Value Theory into a new concept of Dynamic Value. There are two fundamental assertions underlying the development of Dynamic Value: (1) the instantaneous value of a system is the difference between what the stakeholders desire and what the system can deliver at a given point in time, and (2) stakeholders base their decisions on the life cycle value of the system, where life cycle value is defined as the integration of instantaneous value over the period of interest.

[12]The authors actually argue that the field of OSA design is subsumed by their methodology [26].

المنارة للاستشارات
www.manaraa.com

The authors go on to define instantaneous value in commercial design as a purely cost-benefit analysis[13]. To that end, they assume that the Initial Cost & Value (IC&V) desired by the stakeholder is equal to the sum costs of developing, manufacturing, and deploying the baseline system. From that point on, they assume that the stakeholder would only be completely satisfied with an equivalent system composed of SOTA components. Therefore, as the SOTA progresses, the value desired, VD, by the stakeholder increases according to the formulation in Equation 42, where $f_{TA_i}(t)$ represents the change in value driven by technological advances [14]. As implied by the cost-benefit structure, the authors argue that this value desired can be monetized through a conjoint analysis[15]. In addition to the value desired, the authors also note that the stakeholder will likely contended with increasing Maintenance Costs, MC, as the system ages due to wear-out and obsolescence costs. This relationship is provided in Equation 43, where the terms on the right-hand side of the equality reflect wear-out, f_{WC_i} and obsolescence costs, f_{OC_i}, of the i^{th} component [26].

$$VD_i(t) = f_{TA_i}(t) + IC\&V \tag{42}$$

$$MC_i(t) = f_{WC_i}(t) + f_{OC_i}(t) \tag{43}$$

Instantaneous value is then given by the sum of the value desired and the associated maintenance costs. The authors then define the instantaneous Value Loss, VL_i, as the difference between the instantaneous value (i.e. what the stakeholder wants) and the IC&V (i.e. what the system provides). Finally life cycle value loss is given

[13] Recall that cost-benefit monetizes all design considerations, which is a common feature of commercial design literature.

[14] The author's original formulation includes an "Economic Growth" parameter, which is excluded here for the sake of simplicity.

[15] Conjoint analyses are methods to establish consumers' utility for various product attributes, which can then be used to determine how much they would be willing to pay for such an item [36].

www.manaraa.com

by the integral of the instantaneous loss over the system's expected life. This relationship is provided in Equation 44, and depicted in Figure 49 for a static, integrated system [26].

$$VL_i(t) = \int_{t=0}^{T} [VD_i(t) + MC_i(t)]\, dt - IC\&V \tag{44}$$

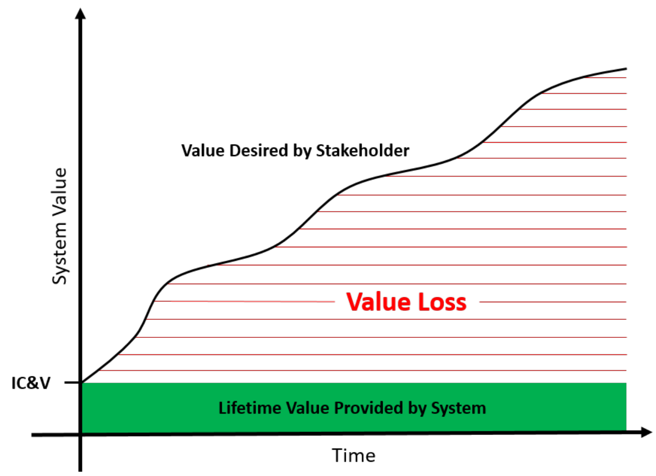


Figure 49: Dynamic Value in the Context of a Closed System

Figure 49 presents a rather bleak picture for system design, but this, the authors argue, can be improved by upgrading the system with SOTA components. When an upgrade occurs, maintenance costs noted in Equation 43 revert back to zero, and the new instantaneous value fully satisfies the stakeholder's desire for the SOTA. This is graphically depicted as a vertical step raising the system value to the value desired by the stakeholder, which in turn reduces the life cycle value loss. Figure 50 depicts this scenario for two competing refresh plans; one in which significant, infrequent upgrades are applied, and one in which less substantial, but more frequent upgrades are applied [26].

As with OSA assumptions, however, these upgrades are not free. The authors idealize the costs of upgrades, $UC_i(t)$, into two categories: development and production cost, DPC_i, and suspension of service costs, SSC_i. This relationship is provided in

المنارة للاستشارات
www.manaraa.com

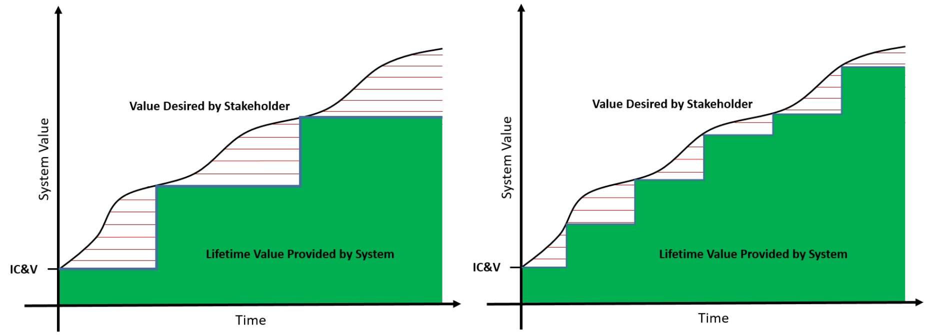


Figure 50: Impact of Alternative Technology Refresh Plans on Dynamic Valuation

Equation 45. At this point, there are two methods under which one could proceed. First, it can be assumed that value loss is a cost with the same units and significance as the cost of upgrades. Under this scenario, the ideal refresh plan simply corresponds to the timing of technology infusions minimizing the total cost of the system in accordance with Equation 46. This is the approach recommended by the authors [26]. If, on the other hand, value loss is not directly comparable to the cost of upgrades, then alternative plans can be found that are efficient with respect to the competing objectives of minimizing value loss and cost. This is an approach that would be more amenable to the OSA context.

$$UC_i(t) = DPC_i(t) + SSC_i(t) \tag{45}$$

$$min = \sum_{i=1}^{n} \left(\int_{t=0}^{T} [VL_i(t) + UC_i(t)]\, dt \right) \tag{46}$$

5.3.3 Establishing the Value of Technology Refresh Plans

Browning and Engel's Dynamic Valuation methodology is entirely predicated on the notion of cost-benefit, and it was well established in Chapter Two that such methods are inappropriate for an acquisitions context. Yet, there is an intuitive appeal in the foundational assertions used to develop the concept of dynamic value. Those

www.manaraa.com

assertions are restated below:

1) The instantaneous value of a system is the difference between what the stakeholders desire and what the system can deliver at a given point in time.

2) Stakeholders base their decisions on the life cycle value of the system, where life cycle value is defined as the integration of instantaneous value over the period of interest.

Consider the application of these assertions to the concept of a performance profile in a deterministic scenario *(uncertainty will be addressed in the next chapter)*. At a given point in time, the stakeholders desire a system that provides maximum effectiveness, subject to the constraint that all performance thresholds are simultaneously satisfied. What the system can deliver is defined by its performance profile, which follows from the refresh planning assumptions governing the timing and selection of technology infusions. A direct application of the fundamental assertions of Dynamic Value would therefore stipulate that the instantaneous value of the system for a particular MoE is given by the difference between the performance and requirement profiles at a single point in time. The second assertion then dictates that the appropriate metric for decision-makers is the total life cycle value of the system, which is determined by integrating instantaneous value over time. This concept can be expressed mathematically by Equation 47, where V_i is the life cycle value for the i^{th} performance measure, $R_i(t)$ is the corresponding system requirement profile, and $P_i(t)$ is the performance profile for a given TRP. In addition, the term δ_i represents the indicator function given in Equation 48, which ensures that any TRP failing to meet performance thresholds is removed from consideration. Figure 51 provides the graphical significance of this formulation to the notional example previously described.

$$V_i = \delta_i \cdot \int_{t=0}^{T} [P_i(t) - R_i(t)]\, dt \qquad (47)$$

المنارة للاستشارات
www.manaraa.com

$$\delta_i = \begin{cases} 1 & \text{if} \quad P_i(t) \geq R_i(t) \; \forall \; t \in [0, T] \\ 0 & \text{otherwise} \end{cases} \tag{48}$$

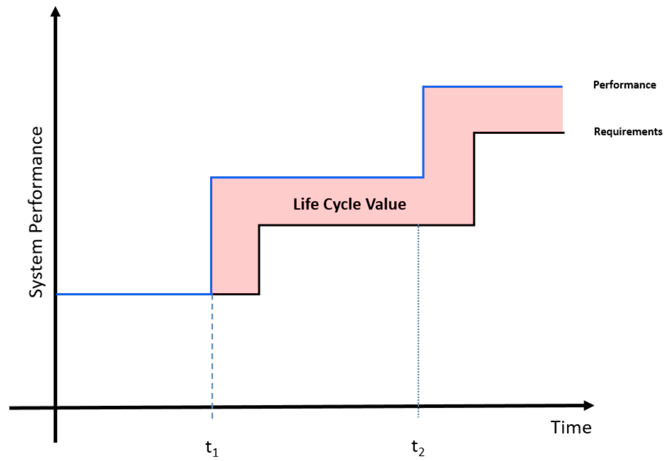


Figure 51: Determination of Dynamic Value for a Notional TRP

Application of the principles of dynamic value requires the decision-maker to look at the problem from an entirely different perspective. From a dimensional standpoint, this new perspective can be thought of altering the unit of measure from units of *performance* to units of *performance·years*. While there maybe some resistance to accepting this way of thinking, the benefit of its application is that alternative TRP's can now be evaluated using the same MODM tools that are currently in use to manage the complexities of static designs. Figure 52 demonstrates why this observation is valid. The left portion of the figure represents an arbitrary number of TRP's that can be formulated by altering assumptions regarding the timing/selection of technology to infuse. The cost of these future upgrades can be reasonably estimated using the cost estimating techniques provided in Chapter 2, since the infusion of an equivalent component with improved attributes is an ideal scenario for the CER and analogy approaches of cost estimation. If one applies the dynamic value transformation to

المنارة للاستشارات
www.manaraa.com

the performance profiles defining these TRP's and plots these results alongside the corresponding cost estimate, then the result is the familiar cost-effectiveness plot. This result is reflected by the dynamic value frontier in the right portion of Figure 52.

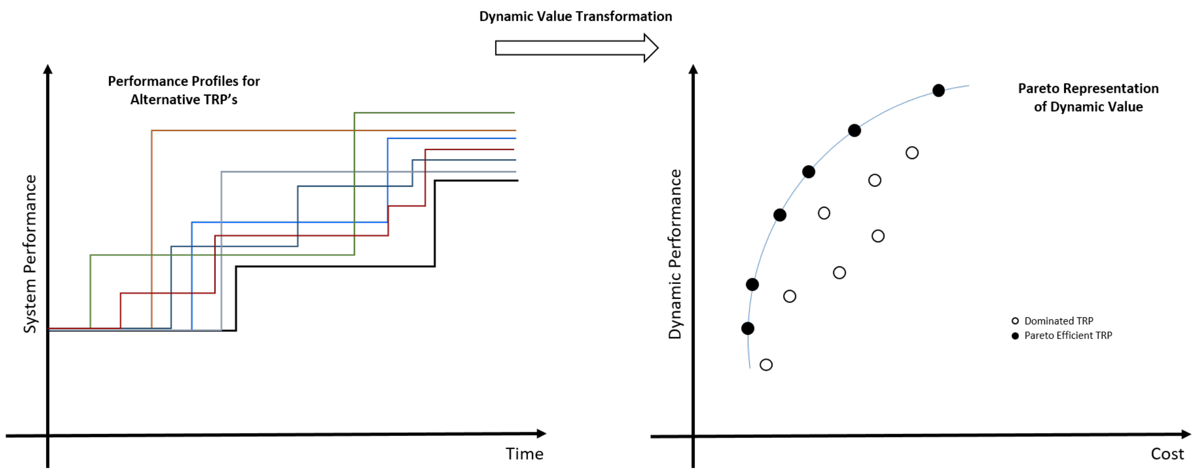


Figure 52: Cost-Effectiveness Representations Under Dynamic Value

Using the representation provided in Figure 52, it is now possible to apply the concept of Pareto efficiency to reduce the dynamic trade space of TRP's. The level of complexity under this representation is equivalent to the complexity present in the trade space exploration for closed architectures, with the caveat that the decision-maker must accept an alternative metric for performance/effectiveness. It is asserted in this dissertation that such a trade space would be manageable for decision-makers, which leads to the conclusion that the evaluation of TRP's in the context of dynamic value satisfies Research Question 2.1.

5.3.4 Alternative Valuation Schemes

The formulation presented in the previous section provides a baseline approach to establish the dynamic value of competing TRP's. However, if one considers the numerous valuation schemes found in Table 10 for traditional Value Theory, then it is clear that there is no exact mechanism to determine an optimal value function. As Value Theory has received considerably greater attention in the literature than

المنارة للاستشارات
www.manaraa.com

Dynamic Value Theory, it follows that the same observation would likely apply. As such, the work of the previous section is not meant to prescribe Equation 47 as the singular formulation that must be used to execute this methodology. Rather, there are many possible formulations that could be applied to achieve the desired end state. It must therefore be left to the analyst to determine, given the context of their problem, what specific mechanism should be applied. The only constraint is that the two fundamental assertions of dynamic value must be respected.

To consider how alternative dynamic valuation strategies could be formulated, consider again the two basic properties of traditional value functions: results are scaled based on their relation to an ideal state, and a weighting mechanism is used to account for the relative importance between performance measures. Neither of these properties impacts the second assertion of dynamic valuation, so the space for creativity is confined to how these properties can be applied to modify the definition of instantaneous value. What follows is a brief set of example formulations intended to aid those interested in applying this methodology.

5.3.4.1 Saturation

The baseline approach to dynamic value implicitly assumes that the ideal state for a system is to maximize performance across all dimensions, at all times. Under this assumption, a closed architecture will invariably have a very high dynamic performance score since it designed to satisfy terminal requirements over the entire life of the system. Yet, there is a distinct possibility that this extreme capacity may not serve a useful purpose early in the system's life cycle. The objective statement provided in Chapter 4 provides an instructive example of why this may prove to be true. In this scenario, the OSA platform requirement was to provide sufficient processing capacity to accommodate improved FMV sensors over time. If tremendous capacity is available from the start, then it is likely that the baseline sensor suite on the platform will

المنارة للاستشارات
www.manaraa.com

not be able to make use of the additional capacity. This implies the existence of a saturation point, where any additional capacity provided beyond this point is irrelevant. If this is true, then the ideal state is defined by a time varying saturation point, as opposed to the terminal requirement. Figure 53 depicts the significance of this assumption. What should be noted in this notional example is that it is now possible for the OSA design to match the dynamic value of a closed architecture system.

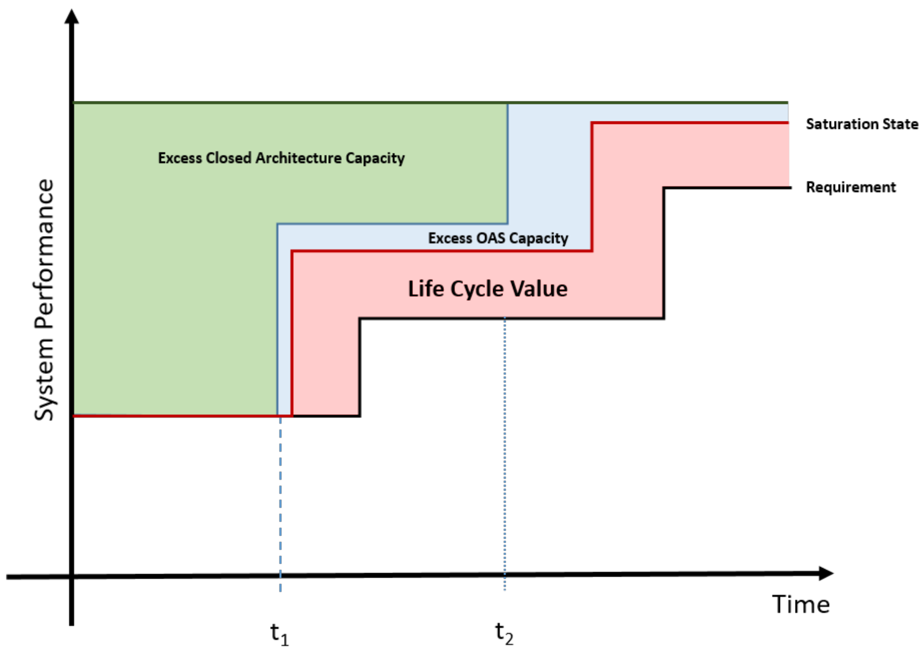


Figure 53: Impact of Saturation State on Dynamic Performance Valuation

While the application of a saturation point may provide added value to the methodology, it also requires the collection of additional information. Unfortunately, this information may be difficult to identify in a complex design analysis, as there are likely many interacting factors that lead to a saturation point. Thus, care should be taken when exercising this approach. With that in mind, if this approach is adopted, then the proper dynamic valuation function is given by Equation 49 below, where the new term, $S_i(t)$ reflects the evolving saturation point as a function of time.

المنارة للاستشارات
www.manaraa.com

$$V_i = \delta_i \cdot \int_{t=0}^{T} \left[min \left\{ S_i(t), P_i(t) \right\} - R_i(t) \right] dt \tag{49}$$

5.3.4.2 Utility Relationships

The use of utility curves/functions is a common methodology used in the DoD to perform trade-off analyses [56]. The purpose of such curves is to capture different relationships defining a decision-maker's belief that there is a non-linear change in value associated with an incremental increase in performance. For example, the relative utility of an attribute can be represented as a constant (straight line), increasing value (concave), or decreasing value (convex), or discrete changes (step) [59]. Figure 54 provides an example of common continuous and discrete forms used in real-world analyses.

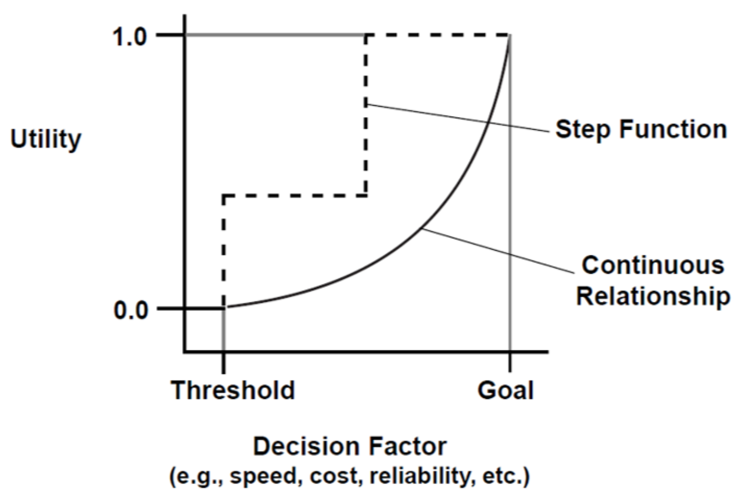


Figure 54: Example Utility Curves [56]

The baseline concept of dynamic value implicitly assumes that a linear relationship exists between an increase in performance and an increase in value. Defining a saturation point provides an upper limit on this increase in value, but the relationship between requirements and saturation is still inherently linear. Thus, applying this concept may provide even greater value to the methodology than either of the

المنارة للاستشارات
www.manaraa.com

previous formulations. As with the saturation scheme, however, application of utility theory to the instantaneous value requires additional information. Specifically, this information must define the utility curve as a function of time over the entire life cycle. While this may appear cumbersome, it may prove useful when paired with the saturation approach, as most utility curves require upper (saturation) and lower (requirement threshold) bounds. In other words, utility theory can be efficiently leveraged if one simply defines a constant functional relationship between the upper and lower bounds. Regardless of the specific application, however, the proper dynamic valuation function is given by Equation 50.

$$V_i = \delta_i \cdot \int_{t=0}^{T} \left[u_i(P_i(t), t) \right] dt \tag{50}$$

5.3.4.3 Satisficing

A common problem in engineering design is formulated as a minimization of the cost of design, subject to constraints that performance simultaneously satisfies a set of minimum thresholds. These problems are collectively referred to as "satisficing" objectives. In a dynamic formulation, this objective must be adapted to account for the variation of requirements and performance over time, but this concept is already enforced via the indicator function in Equation 48. Consequently, if the problem is determined to be a satisficing objective, then the dynamic valuation is simply given by the indicator function. No further modification is required. It should be noted, however, that this is simply the *performance* valuation. The cost of each TRP must still be determined, and an optimization would still be required to identify the TRP with the lowest cost. The approach to structure this optimization has yet to be presented, but will be developed in the next chapter.

المنارة للاستشارات
www.manaraa.com

5.3.5 Review

In conclusion, dynamic valuation provides a means to adequately address Research Question 2.2. This criterion was met by reducing time variation in performance / requirement profiles to a vector of dynamic performance values. Pareto efficiency can then be applied to reduce the objective space to a more concise trade space. Moreover, traditional value theory can be leveraged to aggregate multiple MoE into an aggregate concept of dynamic effectiveness if the decision-makers believe this is warranted. The result is a traditional cost-effectiveness formulation of open systems, where the key distinction is that the abscissa reflects dynamic effectiveness as opposed to static effectiveness. In addition, it is also possible to evaluate closed architectures in terms of dynamic effectiveness, which implies that it is possible to directly compare open and closed architectures. The ability to perform these comparisons is the stated requirement for the fourth step in the proposed methodology, and Figure 55 again updates the process model to record these observations.

Similar to Value Theory, however, there is no academic basis on which to prescribe a *"one-size-fits-all"* approach that will be ideal for all scenarios. Rather, analysts and decision-makers can select whichever formulation suits their needs, so long as the method respects the two fundamental assertions of Dynamic Value Theory.

5.4 Defining the Trade Space of Refresh Plans

The work up to this point established a methodology that proceeds as follows. A MDGM is fit to historical data on a commercial component of interest, which affords a closed form solution for the SOTA curves at future points in time. For a given set of upgrade dates, an analyst then chooses a single point on each of the corresponding frontiers. Each point represents a future component that possesses some combination of improved performance and reduced cost when compared to the previous generation of technology. The improvement in component attributes then propagates through

www.manaraa.com

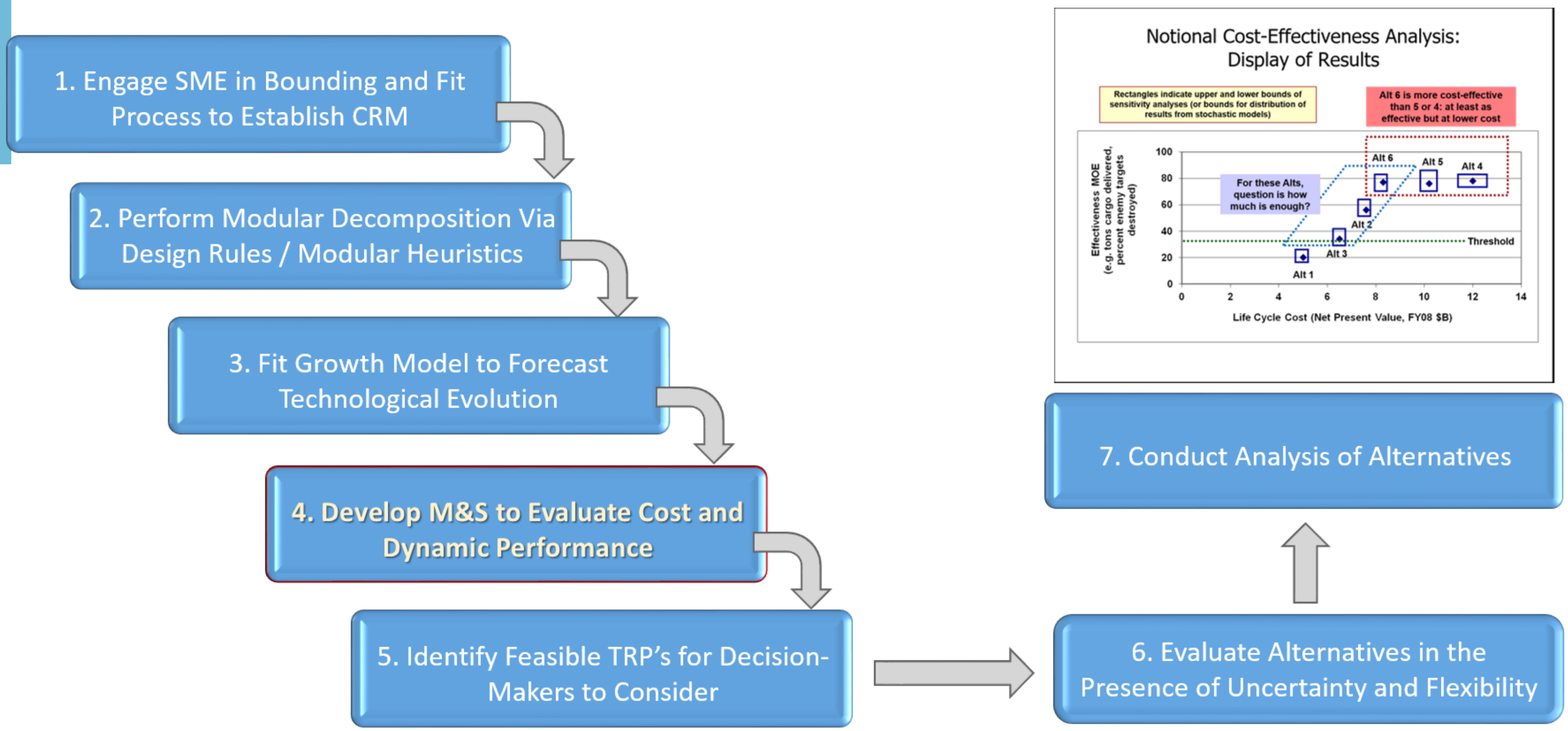


Figure 55: Methodology Update: Step Four - Measures of Performance for Technology Refresh Plans

www.manaraa.com

M&S to generate an improvement in performance attributes. A TRP is defined by the time history of these improvements, which is referred to as the performance profile. A dynamic valuation structure then compares the TRP's performance profile with the system's requirement profile to obtain an n-dimensional vector of responses. Finally, this vector of responses is paired with a cost estimate and compared to other TRP's that were previously evaluated. If this vector is inefficient when compared to the previous TRP data points, then it is omitted from further consideration. If, on the other hand, the current alternative is Pareto efficient, then the option is recorded on a dynamic frontier. This dynamic frontier then provides the trade space in which decision-makers can make intelligent decisions as to whether an OSA is warranted and, if so, how the corresponding TRP should be structured.

The remaining challenge to be addressed in this chapter is how the aforementioned process can be automated in order to ensure that the entirely of the trade space is evaluated. Moreover, this process must be conducted in a highly efficient manner. To highlight the significance of this point, recall that if a development cycle can accommodate T possible upgrades, then there are 2^T unique combinations of upgrade timings. In addition to these alternative timings, any point on the corresponding SOTA frontiers could potentially provide an efficient TRP. If one assumes, as was assumed for the timing of upgrades, that the SOTA curves at future times are discretized into N_c points (recall that each point represents a potential future component) of interest, then the number of unique refresh plans, T_n, is given by Equation 51.

$$T_n = \sum_{i=1}^{T} \binom{T}{i} \cdot \left(N_c^i\right) = \sum_{i=1}^{T} \frac{T!}{i!(T-i)!} \cdot \left(N_c^i\right) \tag{51}$$

Return again to the thought experiment in which there are 20 possible upgrade opportunities, which results in 1,048,576 unique combinations for the timing of infusions. Now, assume that the SOTA curve at each time is evaluated at 20 points uniformly distributed across the frontier. Equation 51 dictates that there are now

المنارة للاستشارات
www.manaraa.com

approximately $2.78 \cdot 10^{26}$ unique TRP's that must be evaluated and filtered to establish the full dynamic frontier. Thus, it is quite clear that a full factorial of evaluations is unrealistic from a computational standpoint. A more efficient approach is required, but it is not immediately clear how that approach should be formulated.

In considering how this question could feasibly be resolved, it was noted that Browning and Engel's original dynamic valuation scheme surely faced the same conundrum. As this method was developed as a valuation mechanism for PPD, it follows that a potential solution may be found within the PPD literature.

5.4.1 Product Platform Design

One of the most significant changes in commercial product design is the recognition that "customers can no longer be lumped together in a huge homogeneous market, but are individuals whose individual wants and needs can be ascertained and fulfilled [140]". As such, commercial companies can gain a tremendous competitive advantage by moving from a single, generic product to a family of products providing the same function, but with attributes closely tailored to smaller cross-sections of the market. This concept is known as the *Mass Customization* paradigm, and the optimization strategies supporting this way of thinking are collectively referred to as Product Platform Design methods.

A product platform can be loosely defined as a modular system architecture, where members of the "product family" members are created by adding, substituting, and/or removing modules [153]. While this approach is conceptually simple, practical implementation has proven to be quite challenging. This is due, in large part, to the difficulty in measuring the costs and benefits of competing designs. In the short term, the initial cost of developing a platform is often much higher than the cost of designing a single product. However, these costs can be diminished by sharing components and production processes across the product family. The goal then is to determine the

المنارة للاستشارات
www.manaraa.com

proper balance of common and unique modules across the product family in order to develop differentiated products efficiently [94].

A flurry of research has been conducted over the last few decades to address this challenge. Simpson identified 40 methods proposed by various authors from 1995 to 2003 alone, which he categorized according to their structural characteristics. These characteristics are enumerated below [153]:

- Module-based product family vs. scale-based family
- Is the platform specified *a priori*?
- Single-objective vs. multi-objective
- Is there a model for manufacturing cost?
- Is there a model for market demand and/or sales?
- Is the formulation deterministic or probabilistic?
- Single stage, dual stage, or multi-stage
- Choice of optimization algorithm

Reviewing these 40 methods and those that have emerged over the last decade is beyond the scope of this thesis. However, juxtaposing these criteria with the OSA design problem previously described allows one to reduce the set of available alternatives. Specifically, an OSA context requires that a method be module-based, multi-objective, specify platform *a posteriori*, and consider uncertainty and manufacturing costs. None of the methods identified in Simpson's survey perfectly satisfies these criteria, but the Multiobjective Genetic Algorithm (MOGA) proposed by Simpson and D'Souza provides a close match[16] [154]. The basic structure of this algorithm is provided in Figure 56.

[16] This method fails to account for uncertainties, but this is the subject of the next chapter.

المنارة للاستشارات
www.manaraa.com

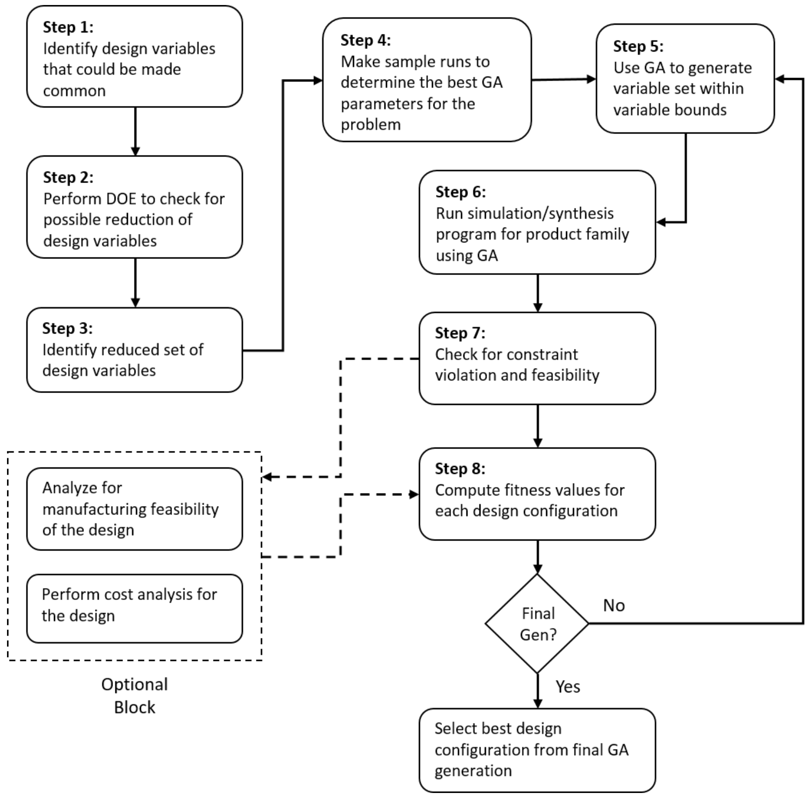


Figure 56: Multiobjective Genetic Algorithm for Product Platform Design [154]

The authors base their methodology on the NSGA-II template. Their ingenuity lies in the nature of the chromosome used to represent families of products. This is achieved by partitioning the chromosome into two elements: a commonality string and a design string. The commonality string has a length equal to the number of design variables, n. Each bit in this string is a true/false representation as to whether that design variable is included in the platform, and thus common across all variants, (true) or unique to each variant (false). The design string is the traditional representation of a system as a vector of attributes, discretized and binned in binary form. There are two caveats to this representation, however. First, each derivative system has its own design sub-string, which means that the length of the platform design chromosome with p derivative systems sub-string has length pn. Second, if a design variable is defined as common in the commonality string, then all instances of of that design variable in the design string are "locked". In other words, one system

المنارة للاستشارات
www.manaraa.com

value is chosen and that value is superimposed on all other design sub-strings.

The NSGA-II algorithm proceeds without modification once this chromosomal representation and the appropriate M&S are established. The second contribution of the author's methodology is the two objective functions used to map alternative product platforms to the objective space. The first objective function operates on the design space using an adaptation of goal programming developed by Mistree et al. [130]. This function, known as the deviation function, is provided in Equation 52, where d_{ij}^{+} and d_{ij}^{-} represent the deviation of the i^{th} DV in the j^{th} system from a predefined technical goal[17]. The second objective function is the Product Family Penalty Function (PFPF) developed by Messac et al. to measure the degree of "commonality" in the family of systems [118]. The PFPF documented in Equation 53 is used to help minimize the relative variation of design variables across all systems, which includes both common and unique variables. The basic assumption is that increasing variation, and therefore decreasing commonality, increases cost. Thus, the classic objective of high performance at low cost is transformed into an objective of low target deviations and low variation [154].

$$Z = \left(\frac{1}{n}\right)\left[\sum_{i=1}^{n}\sum_{j=1}^{p} d_{ij}^{+} + \sum_{i=1}^{n}\sum_{j=1}^{p} d_{ij}^{-}\right] \tag{52}$$

$$\begin{aligned} pvar_i &= \frac{var_j}{\overline{x_j}} \\ var_j &= \sqrt{\left(\frac{1}{p-1}\right)\sum_{i=1}^{p}\left(x_{ij}-\overline{x_j}\right)^2} \quad \text{and} \quad \overline{x_j} = \left(\frac{1}{p}\right)\sum_{i=1}^{p} x_{ij} \end{aligned} \tag{53}$$

The Simpson-D'Souza formulation has some significant limitations. First, the commonality chromosome forces common design variables to be common across all

[17] The plus and minus sign correspond to the "bigger is better" and "smaller is better" quality characteristics.

www.manaraa.com

design variants. This limits exploration of portions of the design space where variables are only common across a subset of variants. In addition, the formulation of the problem as a single chromosome becomes problematic when the scale of the problem, given by the number of variants and/or DV, increases. For instance, Akundi and Simpson evaluated a product family design problem for a family of 10 universal motors with eight DV and three conflicting objectives. They found that application of the previous method took roughly 25,000 generations, with 1,500 members per population, to find an acceptable distribution of efficient points [2].

An advanced method has been proposed by Khajavirad et al. to resolve these challenges [97]. The basic premise underlying this approach is that the best way to make such an algorithm scalable is to restructure the algorithm into a bi-level optimization. In their approach, the top level of the algorithm controls the commonality decisions. This is achieved by converting the single, commonality chromosome into a commonality matrix, with the rows corresponding to product variants and the columns representing design variables. Entries of this matrix take integer values from 1 to p, where any matching values in a column indicate that the design variable is common to those variants. Crossover is achieved by generating two random, integer numbers between $1-p$ and $1-n$, which provide the corresponding row and column indices to partition the matrix into quadrants. One of these quadrants is then randomly selected to be exchanged. In addition, mutations are introduced into the system by considering each entry in the commonality matrix. When a mutation is triggered, the corresponding integer value is changed to toggle the commonality of that DV-product variant pair. These alternative commonality matrices are then evaluated using Equation 54, where u represents the total number of distinct components in the product family, and N_i is defined as the number of distinct integers for the i^{th} DV. This serves as the objective function for the top level of the NSGA-II optimization.

www.manaraa.com

$$CI = 1 - \frac{1}{n(p-1)} \sum_{i=1}^{n} (N_i - 1) \tag{54}$$

The lower level of the optimization process creates one parallel instance for each product in the proposed family. As previously mentioned, any design variables designated to be common must be varied jointly between parallel instances. With that exception, each parallel optimization can run independently to minimize the deviation function provided in Equation 52. The deviations of each variant are then returned to the top-level optimizer and summed. This serves as the second objective function. The efficient results with respect to both objectives are then plotted to provide a two-dimensional trade space for decision-makers to evaluate.

5.4.2 Applicability to Trade Space Exploration

The review of MOGA-based design methods in PPD provides considerable insight into the problem at the beginning of the section. The principal challenge faced by PPD methods is one of combinatorial complexity in the design space. This complexity is derived from the need to determine: (1) which DV should be common among variants, and (2) which design point should be selected for each variant given assumptions of commonality. The similarity of this problem to the problem of trade space exploration for TRP's is readily apparent if one simply recognizes that the issue of commonality is replaced by considerations pertaining to the timing of technology infusions. In other words, the combinatorial complexity for refresh planning is derived from the need to determine: (1) when should technology be infused into the platform, and (2) which design points should be chosen for new components given assumptions of timing. The response developed in the PPD literature was to apply a meta-heuristic approach to efficiently search the trade space. Given the similarity between the two problems, it would stand to reason that this approach would be appropriate for the problem currently under consideration.

المنارة للاستشارات
www.manaraa.com

In addition to the first observation made, the results provided by Akundi and Simpson shine a light on a potential pitfall of a single level formulation that is clearly not self-evident. If the Simpson-D'Souza formulation requires substantial computational resources for modest PPD optimization problems, then it would also stand to reason that the same problem would plague a single level meta-heuristic for modest refresh planning optimizations. Thus, it follows that a two-stage MOGA approach would be the most appropriate formulation for the problem at hand.

The similarity between trade space exploration in OSA design and PPD ends at this point, as the objective functions and two-dimensional crossover/mutation operators are not required for the OSA design problem. With respect to the operators, the basic one-dimensional operators currently in use with the NSGA-II algorithm will prove to be sufficient. In addition, the appropriate objective functions were formulated in the previous section in the form of dynamic performance/effectiveness and TLCC. Thus, the proposed optimization framework to address Research Question 2.3 can be described as follows:

1) A binary string is randomly generated with one character per upgrade opportunity. A "1" turns an upgrade on at that time; and "0" turns the upgrade off.

2) These timings are passed to the next level, which builds the trade space for each point in time where an upgrade occurs by substituting the time parameters into the MDGM provided by Equation 35.

3) A design chromosome is created for each time period; these chromosomes are concatenated to fully define a TRP. This process is repeated to generate a population of candidate TRP's.

4) Each member of the population is evaluated in terms of its dynamic value and TLCC. Results map to non-domination levels and crowding distances.

المنارة للاستشارات
www.manaraa.com

5) The indicator function given in Equation 48 is used to evaluate constraints on performance profiles; non-domination levels are varied accordingly.

6) Check to determine if the NSGA-II algorithm is converged. If not, perform mutation and crossover and return to Step 3.

7) If this is the first iteration, then the current frontier becomes the new trade space; if this is not the first iteration, then the new frontier is superimposed on the existing trade space and inefficient points are excluded.

8) If this is the final timing sequence under consideration, then the process is complete. Otherwise, modify the timing chromosome and repeat Steps 2-7.

This process is also depicted Figure 57, where the inner and outer loops are indicated by the dashed and dotted lines respectively. It is important to note at this point that the algorithm presented here defines the *expected* trade space provided by an OSA. The word expected is intentionally invoked here as this approach does not provide consideration of uncertainty, and is instead conducted with the expected values of all uncertain parameters. As will be demonstrated in the next chapter, introducing uncertainty dramatically increases the complexity and computational burden of any exploration. If one assumes that the decision-makers cannot define *a priori* what their collective preferences are between the different dimensions of dynamic effectiveness and TLCC, then this approach is appropriate. The intent is that this phase of the methodology allows decision-makers to debate where the most cost-effective solutions are on the frontier. The sensitivity of these points to uncertainty and flexibility will then be investigated in the next chapter, but at this point it is also important to note that the process in Figure 57 also addresses the needs of the fifth step in the proposed methodology. Figure 58 updates the process model with this observation.

المنارة للاستشارات
www.manaraa.com

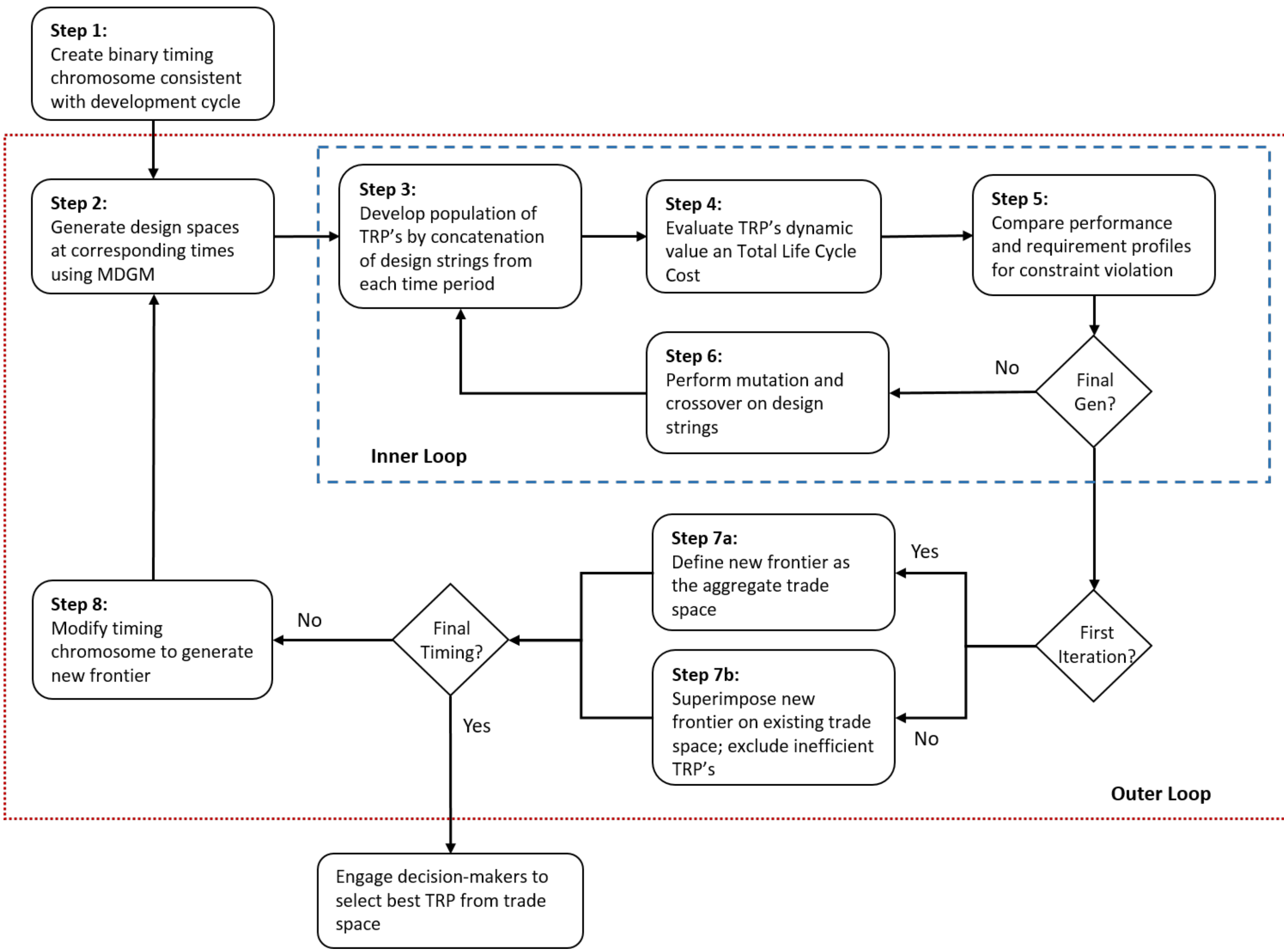


Figure 57: PPD-Insipred Algorithm for Trade Space Exploration of TRP's

www.manaraa.com

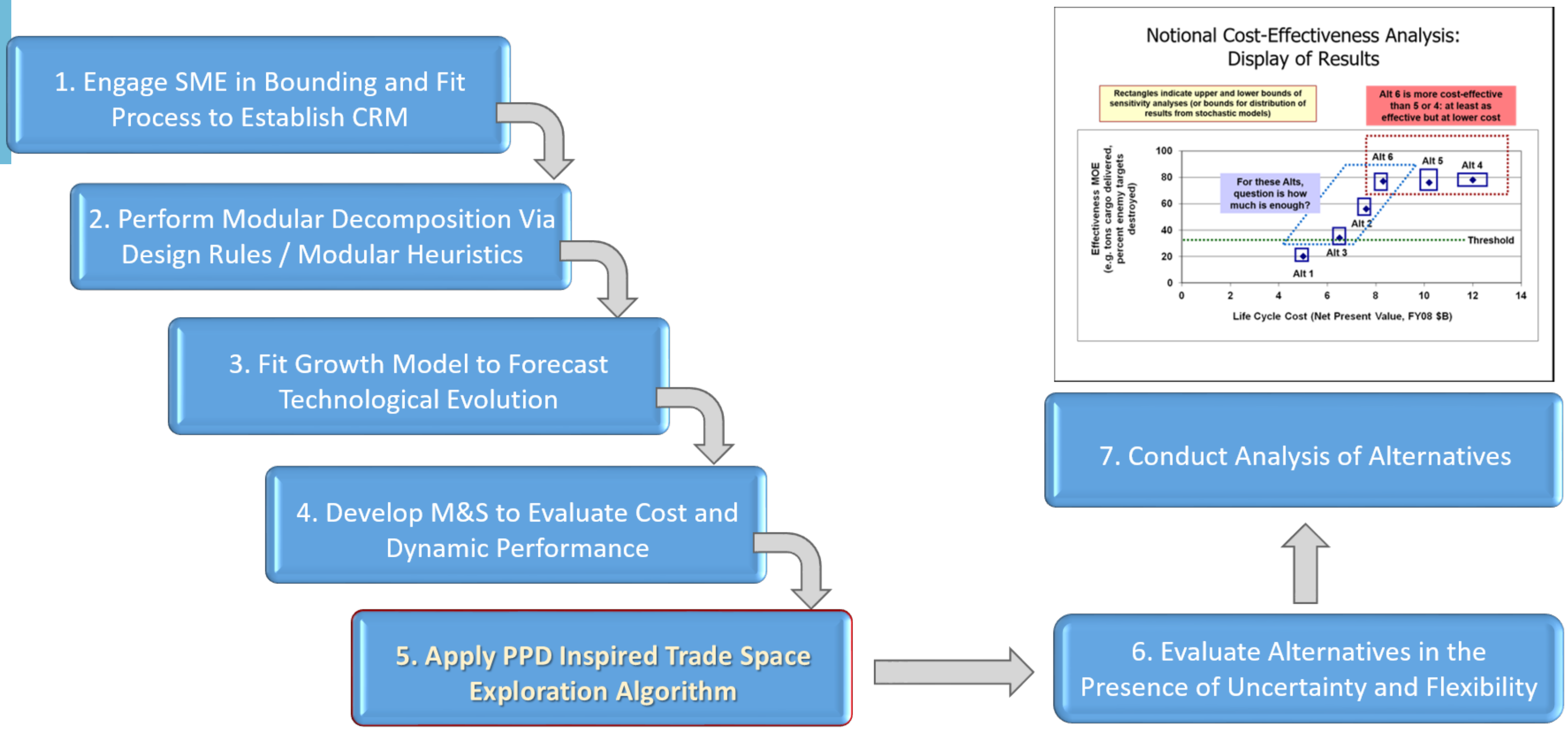


Figure 58: Methodology Update: Step Five - Product Platform Design Inspired Method for Trade Space Exploration

www.manaraa.com

CHAPTER VI

BALANCING UNCERTAINTY AND FLEXIBILITY

The methods developed in the previous chapter provide a means of exploring the trade space available for an OSA platform in terms of the dynamic value and TLCC provided by alternative TRP's. The obvious drawback with this approach is that it is entirely deterministic. In other words, the results generated by the trade space exploration are reliable if, and only if, the future unfolds in exactly the way in which the analysts and Functional SME's expect. One of the main conclusions from Chapter One, however, is that this is highly unlikely, and PM's are therefore reluctant to accept any formulation that includes this assumption. Methods established in Chapter Two (i.e. Risk Assessment and KOSS) attempted to manage this uncertainty through qualitative, SME-driven frameworks. The critique of these methods is that this treatment of uncertainty is too opaque to provide sufficient confidence in the results. Moreover, the implicit assumption in both methods is that a single, rigid development path will be followed, regardless of how the future unfolds. While this assumption is convenient from a modeling perspective, it is fundamentally flawed. One of the great advantages provided by OSA design is the fact that the decision-maker is not committed to a particular development plan. Rather, decision-makers have the flexibility to alter decisions over time in response to new information, and this flexibility serves as a hedge against uncertainty in the co-evolution of requirements and technology.

6.1 Decision Support Methods for Uncertainty in Design

This discussion highlights the fact that a new approach is required to investigate the sensitivity of desirable points in the OSA trade space in the presence of both

المنارة للاستشارات
www.manaraa.com

uncertainty and flexibility. Unfortunately, it is not clear how this new approach should be developed at this stage in the development of the methodology. Therefore, it will be useful to first consider how uncertainty is managed in both MODM formulations and modern PPD methods.

6.1.1 Robust Design

The concept of Robust Design began in the 1980's with Taguchi's "parameter design methodology" in Quality Engineering. This methodology decomposes the design problem into *control factors*[1], $\mathbf{x}$, which can be specified freely by the designer, and *noise factors*, $\mathbf{z}$, that are not under the designer's control. In addition, it is assumed that any variance in the response of a process, $\mathbf{y}$, caused by variations in the noise parameters is governed by an unknown relationship between the noise and control factors. Therefore, the goal is to find the optimal settings of control factors that simultaneously brings the mean of the response to its target value and to minimize the variance around this target. To accomplish this, Taguchi uses a two-part orthogonal array to efficiently structure a design of experiments. These results are then used to determine the settings which maximize the *signal-to-noise* objective function given in Equation 55 [35, 48].

$$S/N = 10log_{10}\left[\frac{1}{n}\sum_{i=1}^{n}(y_i - \overline{y})^2\right] \tag{55}$$

Though Taguchi is given credit for establishing the field of Robust Design, it is well established that these methods are inaccurate for design problems with highly non-linear behavior [35]. Modern Robust Design Methods take a slightly different approach by removing the concept of a noise variable and assuming that the variation in responses stems directly from uncertainty in the DV. Consequently, total ordering of the design space functions becomes problematic, as the propagation of this uncertainty

[1] Control factors are equivalent to the DV previously defined.

www.manaraa.com

through M&S converts the output of the value function from a scalar to a distribution of potential values. Most modern methods typically account for this added degree of freedom by specifying that the value function blend the desire to simultaneously minimize the mean and the variance of the response. Equation 56 gives an example of such a method [34], known as Compromise Programming, where μ_f^* and σ_f^* are the mean and standard deviation at the utopia point, and ϵ is a scaling factor.

$$\begin{aligned} &\text{minimize } \beta \\ &\text{subject to} \quad w_1\left(\frac{\mu_f}{\mu_f^*} - 1.0 + \epsilon_1\right) \leq \beta \\ &\qquad\qquad\quad\; w_2\left(\frac{\sigma_f}{\sigma_f^*} - 1.0 + \epsilon_2\right) \leq \beta \end{aligned} \tag{56}$$

An interesting aspect of these methods is that they are generally structured as a nested loop. The top level is the classic optimizer that searches the objective space for the minimum output of the value function[2], but this evaluation requires the mean and variance of responses at that point. The lower level determines these parameters by either tabulating statistics from a Monte Carlo Simulation (MCS), or using the linearized formulations provided in Equation 57. The consequence of this approach is that the optimizer will prefer flatter spaces in the objective space to sharp peaks, even though the optimal, deterministic value may lie at such a peak. Figure 59 depicts why such a formulation is advantageous.

$$\begin{aligned} &\text{Response Mean} \quad && \mu_y = f(\mu_{\mathbf{x}}) \\ &\text{Response Variance} \quad && \sigma_y^2 = \sum_{i=1}^{k} \left[\frac{\delta f}{\delta x_i}\right]_{x=\mu_{x_i}} \sigma_{x_i}^2 \end{aligned} \tag{57}$$

[2]Robust Design methods are generally structured as a minimization problem because variance is always minimized.

المنارة للاستشارات
www.manaraa.com

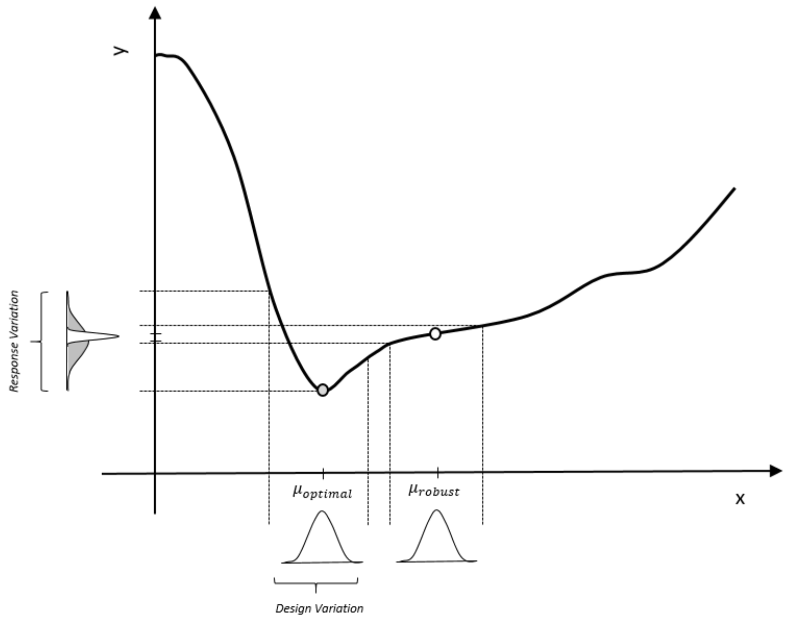


Figure 59: Notional Application of Robust Design

6.1.2 Reliability Based Design Optimization

Robust Design methods based on signal-to-noise ratios or similar constructs are predicated on minimizing the sensitivity of a resulting design to uncertainty in DV. This approach, however, does not explicitly provide confidence that the resulting design will be capable of satisfying requirements, especially when those requirements are active at the optimum. These concerns can be directly addressed by applying the methods from a different class of optimization algorithms - Reliability Based Design Optimization (RBDO). In RBDO, the general formulation consists of minimizing a deterministic objective function[3] in the presence of probabilistic constraints. This formulation is expressed by Equation 58, where $\theta \in \Re^n$ is the vector of random design variables and G_i is the i^{th} constraint that, by convention, is violated when $G_i < 0$. As indicated, the uncertain nature of these constraints dictates that a design is acceptable if, and only if, the probability of a constraint violation, P_{f_i}, is less than or

[3] Although the objective function typically contains random variables, most RBDO consider it as a deterministic value by calculating it with the mean of any uncertain parameters [109].

المنارة للاستشارات
www.manaraa.com

equal an acceptable probability of failure for all constraints, $P_{f_i}^{allowable}$ [1].

$$\begin{aligned} &\text{minimize} \quad V(\mathbf{x}, \theta) \\ &\text{subject to} \quad P_{f_i} = P(G_i\left[\mathbf{x}, \theta\right) \leq 0] \leq P_{f_i}^{allowable} \qquad \text{for } i = 1, \ldots, m \end{aligned} \tag{58}$$

The primary challenge in RBDO is how one evaluates the constraints. The exact value of P_{f_i} for each constraint can be determined by evaluating the integral in Equation 59, where $f_\theta(\theta)$ is the joint Probability Density Function of the random variables. In general, only the marginal distributions and correlation coefficients of the individual random variables are known, which implies that the joint PDF can only be ascertained if all random variables are mutually independent. If this assumption is invalid, which is often the case, then the evaluation of Equation 59 is nearly impossible [109].

$$P(G_i\left[\mathbf{x}, \theta\right] \leq 0) = \int \ldots \int_{G_i(\mathbf{x}, \theta) \leq 0)} f_\theta(\theta) d\theta \tag{59}$$

To resolve this challenge, most RBDO methods transform the integral in 59 from its native space, θ, to a standard Gaussian space, $\mathbf{U}$, through a *Rosenblatt Transformation*. As depicted in Figure 60, the principal benefit of this transformation is that iso-probability levels in the U space are circular, and centered on the origin. The Most Probable Point (MPP) of failure is therefore the point on $G_i(\mathbf{x}, \theta = 0)$, referred to as the limit state, closest to the origin. If the magnitude of the vector from the origin to the MPP is defined by β, then it follows that $P_{f_i} \approx \Phi(-\beta_i)$ and $P_{f_i}^{allowable} \approx \Phi(-\beta_i^{target})$, where Φ is the standard, Gaussian CDF and $beta_i^{target}$ is the transformed target for reliability [170, 109].

With this in mind, the major distinction between alternative RBDO methods is how this optimization process is executed. Different methods exist to determine the MPP, with the most common being the Reliability Index Approach (RIA) and

www.manaraa.com

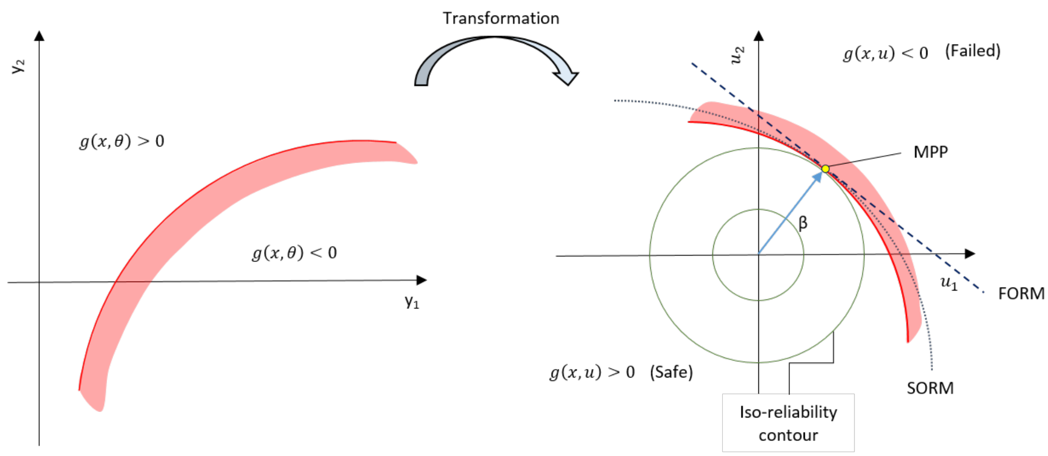


Figure 60: Reliability Analysis [1]

the Performance Measure Approach (PMA). Other methods model the limit state in the neighborhood of the MPP point with either a linear or quadratic surface, which are referred to, respectively as First Order Methods (FORM) and Second Order Methods (SORM). Finally, the two segments of RBDO analyses, optimizing DV's and evaluating probabilistic constraints, can be structured as a two-level, nested optimization, a single optimization, or a sequential optimization. The interested reader is referred to Aoues and Chateauneuf for a more detailed discussion on the precise formulation of these alternative methods [8].

6.1.3 Probabilistic Design

A possible pitfall of RBDO is that a given design point with substantially better reliability, but negligibly worse performance than an "optimal" point found in RDBO algorithms will likely be excluded. The reason for this exclusion is that reliability is used a constraint, as opposed to a value driver to be included in the objective function [48]. This issue is explicitly dealt with through Probabilistic Design Methods, which explicitly treat reliability as an RV. An example of such a technique is Robust Design Simulation (RDS) developed by Mavris et al. [122]. RDS functions in a similar manner as the bi-level RBDO algorithms in that the top level optimizer searches the

المنارة للاستشارات
www.manaraa.com

design space of efficient points, while the lower level manages the reliability analysis. The distinction in the RDS approach lies in the fact that the reliability analysis approximates the locus of MPP for a finite set of reliability levels using an adaptation of the mean value linearization provided in equation 60. By finding the design points which maximize the various reliability levels, RDS is able to return a CDF of possible objective values vs. the risk associated with attaining those levels. Decision-makers can then use these CDF's to make an informed decision of risk vs. reward associated pursuing different design concepts.

$$f(\mathbf{X}) = f(\mu) + \sum_{i=1}^{n} \left[\frac{\delta f}{\delta x_i} \right]_{x_i=\mu_i} (x_i - \mu_i) \tag{60}$$

RDS was later augmented to account for the joint cumulative distributions of multiple objectives in the Joint Probabilistic Decision Making (JPDM) approach [16]. JPDM takes the inverse approach to RBDO by assuming that constraints are deterministic, but the DV, and by extension the RV, are stochastic. This method applies the now familiar bi-level approach, where the top level manipulates the mean of the DV, and the lower level performs an MCS to generate a joint distribution of RV[4]. This joint distribution is then superimposed on the deterministic constraints in order to determine the probability that a design defined by the vector of mean values will satisfy deterministic constraints. This scalar quantity is referred to as the Probability of Success (POS), which serves as the objective function that the top level optimizer seeks to maximize. Figure 61 graphically depicts an application of JPDM to a notional design problem.

6.1.4 Review of Methods and Gap Analysis

One common aspect across across all methods developed to accommodate uncertainty is that each method aggregates RV into a single measure of effectiveness, and then

[4]This normally time-consuming process is expedited through the use of a regression model to approximate the output of detailed M&S.

www.manaraa.com

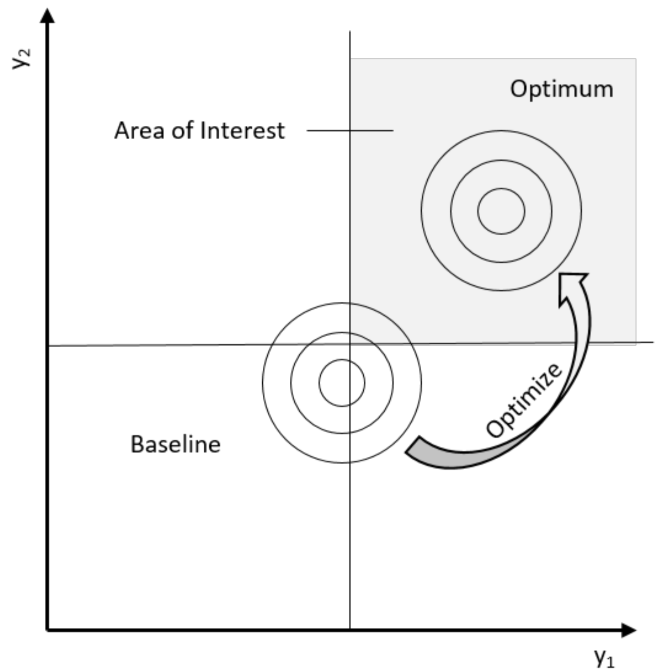


Figure 61: Joint Probabilistic Decision Making [16]

performs a univariate optimization on this measure. This is an important observation. In order to take advantage of any of the methods described in the previous section, it will be necessary to find a similar measure of value for OSA design concepts.

To develop this measure, consider again the discussion surrounding partial ordering of the design space under a Pareto efficiency scheme. The reason for this approach is the assumption that it is often difficult, if not impossible, for a group of decision-makers to agree on their preferences between competing dimensions of dynamic value and TLCC *a priori*. Consequently, decision-makers need to be able to consider the available trade space before any decisions can be made as to which design points should be investigated further. Once these design points have been chosen, however, decision-makers have effectively specified their preferences between competing attributes *a posteriori*. In other words, if an analyst selects an appropriate value function, such as those documented in Table 10, then it is possible to find the set of weights, $\widehat{\mathbf{w}}$, maximizing the objective function at the chosen point, $\widehat{\mathbf{y}}$, using the

www.manaraa.com

simple optimization in Equation 61.

$$
\begin{aligned}
\max \quad & V(\mathbf{w}, \mathbf{y}) = \sum_{i=1}^{m} w_i \cdot \nu_i(y_i) \\
\text{s.t.} \quad & \sum_{i=1}^{m} w_i \;=\; 1 \\
& \mathbf{y} = \widehat{\mathbf{y}}
\end{aligned}
\tag{61}
$$

There is an alternative way to view the optimization posed by Equation 61. Under most value functions, the individual weight parameters dictate the slope of the iso-value curves. Therefore, the optimal weight settings are those that cause the iso-value curves to be tangent to the trade surface at the design point identified by the decision makers. This is depicted in Figure 62. In addition, it is important to note that the decision-makers are not restricted to a single point on the trade space that is worthy of further consideration. Multiple points can be selected, each resulting in an alternative weighting scheme.

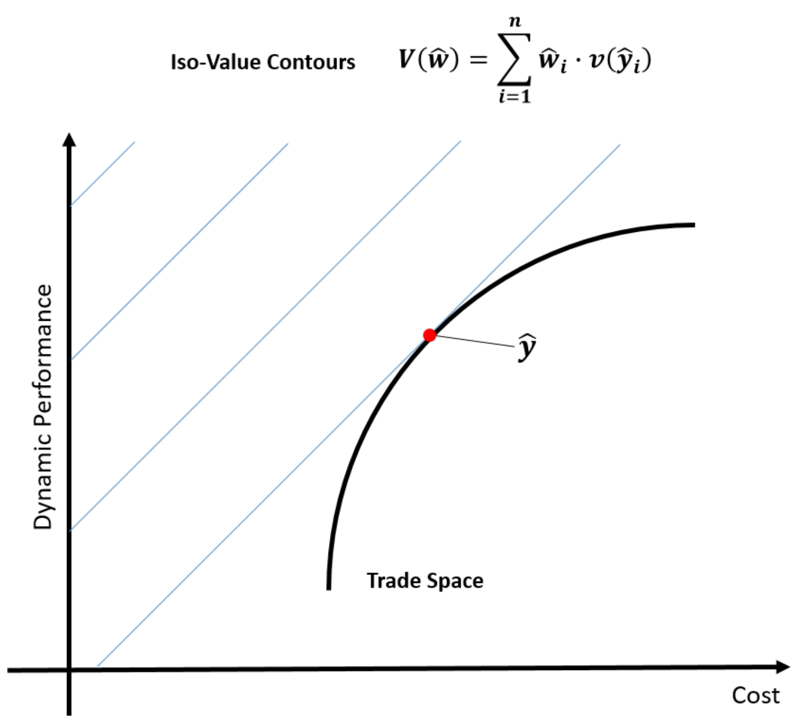


Figure 62: Relationship between Ideal Weight Scheme and Efficient Design Point

www.manaraa.com

This formulation is necessary because the trade space exploration process described in Figure 57 is computationally intensive in its native form. If one were to integrate a form of MCS into the MOGA approach, then the resulting computational burden would increase by several orders of magnitude. Moreover, it is not clear how the trade space exploration process could be altered to account for this uncertainty. In the presence of uncertainty, each iteration of the MOGA would create a new Pareto frontier that would be consistent with the version of future events created by a particular run of the MCS. To date, the author could find no solution in the academic literature that allows probabilistic frontiers to be consolidated into a single trade space visualization. In light of these observations, this approach is deemed to be necessary.

With this in mind, use of the appropriate value function would allow any of the decision support methods for uncertainty described in the previous section to be applied to the OSA design problem. It must be noted, however, that none of these methods consider time variation in the objective space. If the objective space is static, then there is no room for decision -makers to use their flexibility to alter prior decisions as a hedge against uncertainty. Thus, while construction of a value function mirroring the decision-maker's *a posteriori* preferences will prove to be a necessary step, it is not sufficient. The next section will consider methods from other domains that deal with similar challenges in an effort to find an additional mechanism to close this gap.

6.2 Alternative Decision Support Methods

Given that the classic decision support methods found in MODM literature do not provide explicit consideration of time variation in the objective space, it is appropriate to extend the search into adjacent domains. Given the success with finding solutions within the PPD literature, the methods for managing uncertainty in this domain

المنارة للاستشارات
www.manaraa.com

would appear to be a reasonable starting point.

6.2.1 Management of Uncertainty in PPD Literature

As previously noted, system diversity provides value by reaching a larger market, but this diversity can be costly. The previous methods model these costs and benefits indirectly, using commonality and technical performance as surrogates for cost and revenue. A second class of methods attempts to model costs and benefits directly. These parameters are then combined with market uncertainties into a single economic criterion in order to determine which combinations of DV and modules maximize the total value to the company.

The traditional criterion for commercial ventures is the Net Present Value (NPV) approach provided in Equation 62 [29]. Uncertainty in this model is typically dealt with by either adding a "risk premium" to the discount factor, i, or calculating the expected value of uncertainty in future design requirements, technology, and market conditions. However, recent research in product platform valuation argues that this approach is inappropriate for two reasons: (1) the value of a platform strategy does not depend linearly on these uncertainties, and (2) the deployment of future variants is not mandatory, but at the discretion of decision-makers using information available to them at the time [94]. A new approach was therefore required.

$$NPV = \sum_{t=1}^{N} \frac{F_t}{(1+i)^t} \tag{62}$$

Where:

- i = Effective interest rate per period
- N = Number of compounding periods
- F_t = Future sum of money at time t

المنارة للاستشارات
www.manaraa.com

The design and deployment of product variants under uncertain conditions has been the subject of extensive research. Much of this research has coalesced around the concept of Real Options as an alternative to traditional NPV valuation [94]. One of the most popular methods in this category is the methodology proposed by Gonzalez-Zugasti et al. [72]. This formulation applies a two-step approach. In the first step, the platform is selected and frozen using the trade space exploration process previously described. In the second step, future variants are determined through consideration of both uncertainty that cannot be reduced, and the flexibility to adapt actions in response to new information. The conceptual framework for this later step is the decision tree structure depicted in Figure 63.

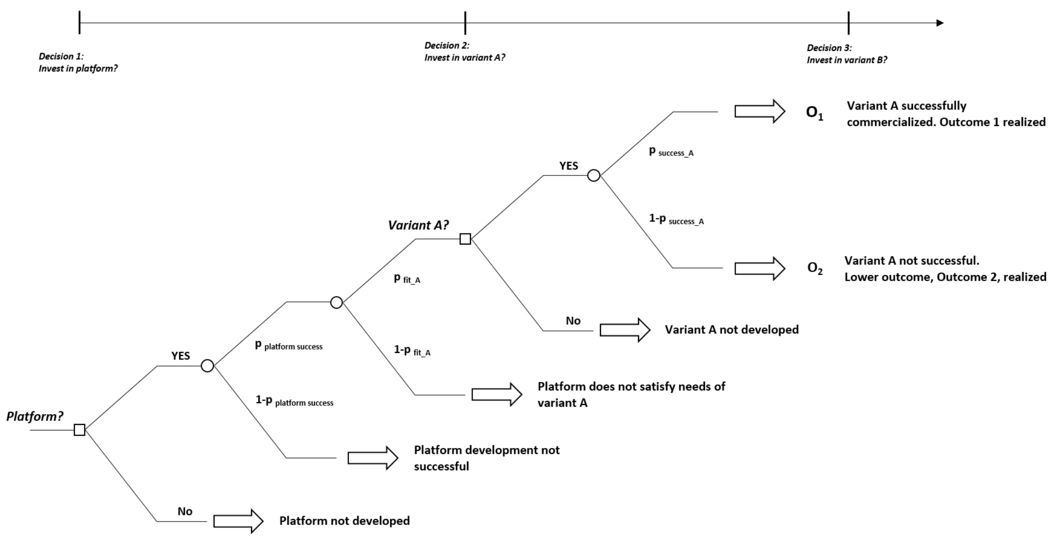


Figure 63: Modeling Uncertainty for Product Platform Evaluation [72]

Developing this tree structure requires enumerating the failure points along the development path of both the product and the platform, along with the probability of these events occurring. These failure points are indicated by the square nodes of the tree. In addition, a second of set points, indicated by circular nodes, are superimposed on the tree where decision-makers can influence the process by either continuing with, or discontinuing the development. Finally, the outcomes for all terminal branches of the tree are estimated. The valuation scheme for the completed variant tree starts

www.manaraa.com

at the latest square node (i.e. the farthest to the right). The value of the process at this point is given by Equation 63. Note that the first term is the difference between the expected value of the process at the following circular node, which is beyond the control of decision-makers, and the investment required to obtain those possible outcomes, I_A. The maximum of this value and zero reflects the decision-makers ability to terminate the program at that point, which is the principal distinction between the Real Options approach and the Expected Value approach. This process then proceeds recursively from right to left until the base node is reached. The value at this base node then represents the value of the platform.

$$V = max\left\{p \cdot O_1 + (1-p) \cdot O_2 - I_A, 0\right\} \tag{63}$$

Though the example in Figure 63 represents a fairly simple case, the same process can be used to model complex options across an arbitrarily large set of variant options. While this is a desirable property, the drawback of this approach is its sensitivity to the probabilities of various events occurring. It is not clear how one can objectively amalgamate the various sources of uncertainty into a single measure of probability. Moreover, the authors do not describe a mechanism to integrate the joint probabilities of more complex scenarios.

These complicating factors exist for any generic system concept, but they are particularly problematic for an OSA design context. In particular, there is no clear way in which the various drivers of uncertainty in the evolution of requirements, technological development, modeling assumptions, etc. could be adapted into a single probability measure. In addition, the PPD concept lends itself well to the finite lattice structure given in Figure 63, but the number of scenarios one could contemplate for an OSA would constitute an infinite set[5]. No tree structure could be developed to efficiently

[5]This statement reflects the understanding that uncertainty in the joint-evolution in requirements, technology, etc. is modeled as a continuous, rather than a discrete, process.

المنارة للاستشارات
www.manaraa.com

capture this complexity. Moreover, the tree structure depends on the analyst's ability to estimate the terminal outcomes of each branch of the tree. Therefore, even if it were possible to formulate a decision tree capable of adequately managing the complexity of uncertainty, it is not clear how these outcomes could be estimated with any degree of consistency and objectivity. Finally, the optimal decisions a decision-maker should make in response to new information are based on the joint probabilities of those factors driving events. The ability to model these interactions is therefore a mandatory requirement for the problem at hand.

In review, the popular PPD method for accounting for a decision-maker's flexibility serves as a useful starting point to identify requirements that this methodology should accommodate. With that in mind, however, it is self-evident that this particular approach cannot be reconfigured to suit the needs of the problem statement in Research Question 3. Given that the inspiration for the methodology proposed by Gonzalez-Zugasti et al. came from the field of Real Options, that would seem to be a likely domain in which a more suitable approach can be found.

6.2.2 Real Options

The field of Real Options builds on an analogy between the options available to designers in engineering projects and financial option contracts [24, 134]. The latter are contracts that enable the holder of the option to buy or sell an asset at some point in the future, but without any obligation to do so if the net outcome were to be negative. The proper method to attach a dollar value to these contracts was an unsolved problem since the early 20th century [94]. An elegant solution to this riddle was eventually found in 1973 through the famed "Black-Scholes equation", which provides a closed form solution to determine the dollar value of options that allow an asset to be purchased on a specific date in the future [20]. Such a contract is referred to as a European Call option, where the "European" title implies the contract can

المنارة للاستشارات
www.manaraa.com

only be exercised at a specific time, and the "Call" qualifier indicates that the buyer has the right to purchase, rather than sell, the underlying asset at a given price.

Hundreds of publications have since used the same or a similar construction to value options with a variety of characteristics. Among these publications is the highly popular lattice-based method developed by Cox et al. for the valuation of options that can be exercised at any time prior to expiration (American options) [44]. In addition, a litany of competing approaches have been developed to model options with a discrete set of exercise dates (Bermudan options) [92, 108] and/or multiple exercise opportunities (Exotic options) [87, 86]. This deluge of research has even extended beyond the range of valuation for financial instruments to consider the added value of having the "option" to take some action at a later date in the design process. These options pertaining to physical systems, as opposed to digital currency conversions, are collectively referred to as Real Options, which Neufville defines as follows [134]:

> "Real" options deal with physical things rather than financial contracts. Specifically, they refer to elements of a system that provide "rights, not obligations" to achieve some goal or activity. Generally speaking, all elements of a system that provide flexibility can be considered as "real options".

This notion that flexibility in design can be modeled using a real options framework has gained substantial traction in recent years. Examples include the phased deployment of communication satellite constellations under demand uncertainty [49], decisions on component commonality between aircraft in the same product family [114], and building design under rent and space utilization uncertainty [94]. In addition to this raw academic inertia, there is a strong intuitive appeal in the use of real options in OSA design. Specifically, there is no obligation to upgrade at a given point, much

المنارة للاستشارات
www.manaraa.com

less an obligation to upgrade with a specific component. OSA design principles therefore provide the right, without the corresponding obligation, to increase the dynamic value of the system in exchange for an increase in TLCC. Moreover, this "option" to upgrade provides flexibility to decision-makers to adjust previous decisions in response to new information. It would therefore appear that OSA fit all of the criteria to be successfully modeled within the scope of real options theories and methods.

There is, however, a significant challenge in the practical application of the previous assertions - all real-options-related research is structured according to a cost-benefit framework. In other words, application of any real-options-based method requires all of the value drivers in the OSA to be monetized and aggregated, similar to the Baldwin and Clark approach [15], into a net value. The various real and financial options frameworks then provide the means to maximize the monetary return on investment. As was demonstrated conclusively in Chapter 2, this is simply not possible. The realities of the acquisitions context of design dictate that value drivers defy any form of monetization, and are generally incomparable with one another. Thus, there is no consistent, objective means of reducing the OSA design problem into a single, monetary value. The conclusion then is that no existing real options framework will be sufficient for the problem at hand.

These observations may appear daunting, but there is an emerging consensus among some academics that options theory can actually be decomposed into two categories. Steffans and Douglas describe this dichotomy as follows [158]:

> Real option approaches have been posited as both an analytical tool to value specific opportunities, that is Real Options Valuation (ROV), and as a strategic heuristic to aid decision making under conditions of uncertainty, often referred to as "real options thinking".

The authors argue that the best approach in the presence of competing objectives and design flexibility is to use the heuristics associated with "real options thinking",

www.manaraa.com

but forget the real options valuation methods. This advice may prove well-founded for the problem of blending uncertainty and flexibility in OSA design. In considering what heuristics may prove fruitful to the problem at hand, it is useful to first screen out the categories of methods that are clearly inappropriate for the acquisitions context. With this in mind, the previous discussion highlighted that options methods are commonly grouped according to the number and timing of potential exercise opportunities: European, American, Bermudan, and Exotic. Given that exercise opportunities equate to upgrade opportunities, it follows that both American and European options can be removed from consideration since they are structured to accommodate only a single exercise *(i.e. upgrade)* date. This leaves only Bermudan and Exotic option methods as the remaining alternatives from which one could derive useful heuristics.

While this down selection is useful, there are still a multitude of authors who have contributed methodologies to this category over the years. This includes [94]: Longstaff and Schwartz (2001)[108], Andersen (2000)[5], Ibanez and Zapatero (2004)[87], Barraquand and Matineau (1995)[17], and Rayman and Zwecher (1997)[144]. A complete review of these methods is beyond the scope of this dissertation; however there is one clear theme across all of the methods that were reviewed within this category - valuation of the options is determined by simulation.

Any valuation conducted via simulation requires the use of a growth model. By far the most common model, which is the same model used in the formulation of the Black-Scholes equation, is the stochastic differential equation of Geometric Brownian Motion (GBM) provided in Equation 64. This model is interpreted as follows: a relative change in stock value, S_t, is a combination of a deterministic proportional growth term, μdt, similar to inflation or interest rate growth, and a normally distributed fluctuation, σdW_t, where W_t is standard Brownian motion [61]. Using this model, it is possible to simulate alternative versions of how the stock price, which

المنارة للاستشارات
www.manaraa.com

is the underlying asset in financial option frameworks, will vary over time through a standard MCS. Figure 64 provides an example of this form of simulation, and Figure 65 depicts the resulting time variation in the PDF of stock price by evaluating a cross section of these sample paths at specific points in time.

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t \tag{64}$$

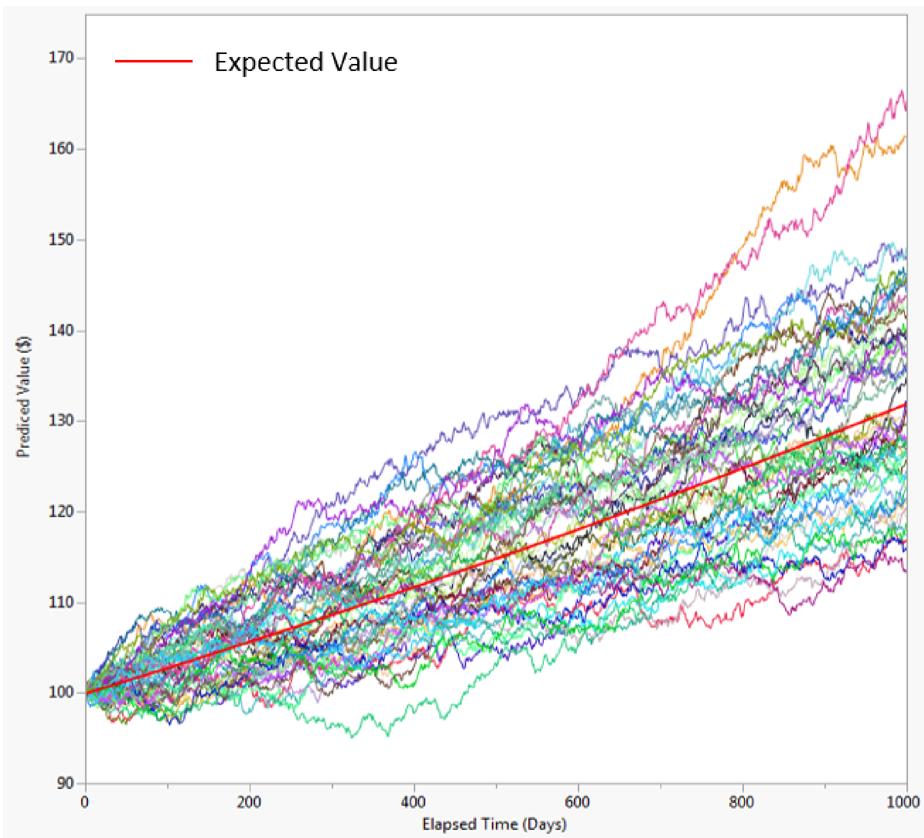


Figure 64: Monte Carlo Simulation of Stock Prices Under GBM

Given the simulations generated by an underlying growth model, each of the methods reviewed in the category of options methods applies a similar heuristic that can be summarized as follows. *Each iteration of the MCS generates an alternative version of future events that is consistent with the present day beliefs regarding uncertainty. Given that the the future is known under the scenario generated via the MCS, an*

المنارة للاستشارات
www.manaraa.com

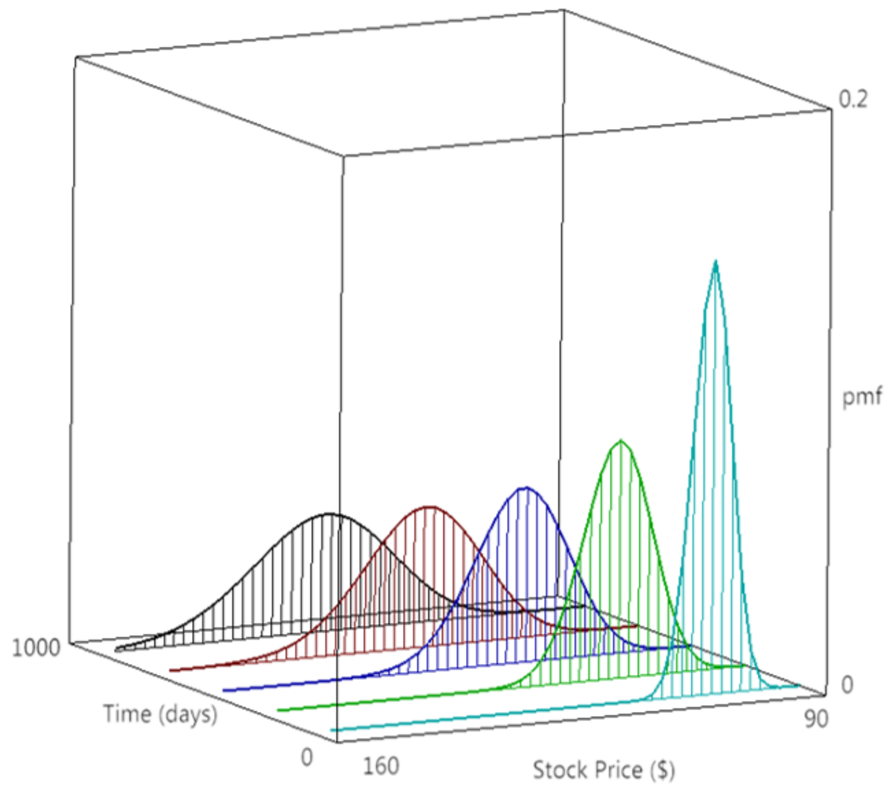


Figure 65: Time Variation in PDF of Stock Price over Time

algorithm can determine the optimal course of action that should be pursued. The results of each iteration are recorded and, given that each MCS is deemed to be equally likely, the average result across all optimal courses of action is believed to be the true value.

That being said, most of the methods of interest in this review go on to provide more complicated formulations to extract the optimal decision-making strategy that a trader can use for real-time decision making. While this added feature would certainly be desirable for the OSA methodology under development, it is not clear how this optimal exercise strategy could be adopted to a cost-effectiveness formulation. This feature will therefore be left for further refinements of the methodology, but the core heuristic is directly applicable to the work established up to this point. The next section will demonstrate this point.

المنارة للاستشارات
www.manaraa.com

6.2.3 Simultaneous Modeling of Uncertainty and Flexibility

The discussion in the previous section determined that an appropriate method to simultaneously model both uncertainty and the moderating effects of flexibility can be found by direct application of the real options "pricing by simulation" heuristic. This can be restated as follows:

> Uncertainty and flexibility can be simultaneously enforced by simulating alternative versions of future events, and imposing the optimal course of action across all simulations. Since each simulation is equally likely, the distributions of these results reflect the true likelihood of outcomes.

Section 6.1 determined that the "optimal" course of action in a given simulation is the TRP that maximizes the value function corresponding to the design point in the expected trade space that was chosen by decision-makers. With this in mind, it is possible to specify a continuous, bi-level optimization statement for each iteration of the MCS as follows. Assume that the same number of upgrades used to generate the initial design point on the trade space are applied to the MCS search algorithm. Let the top level optimizer specify the timing of each upgrade, and pass these timings to the lower level optimizer. The lower level optimizer constructs the design spaces available at each time with a MDGM, and searches over these design spaces to find the point that maximizes the given value function. This result is then returned to the top level optimizer. If the results converge, then optimization ends, and the results are recorded. Otherwise, the timing of upgrades is perturbed, and the inner loop iterates. This process is depicted as a flow chart in Figure 67. In addition, the optimization statement that would be applied for an instance in which two upgrades are perused is provided in Equation 65 (top level) and Equation 66 (lower level).

www.manaraa.com

$$\begin{aligned} \min \quad & -V\left[y(\alpha, \beta, x)\right] \\ \text{subject to} \quad & \alpha > 0 \\ & \beta - \alpha > 0 \\ & 1 - \beta > 0 \end{aligned} \tag{65}$$

$$\begin{aligned} \min \quad & -V\left[y(\alpha, \beta, x)\right] \\ \text{subject to} \quad & \delta_i > 0 \qquad\qquad\qquad \forall\, i = 1, \dots, m \\ & x_{i,0} \;<\; x_i^{t=\alpha} \;< L_i \quad \forall\, i = 1, \dots, n \\ & x_{i,0} \;<\; x_i^{t=\beta} \;< L_i \quad \forall\, i = 1, \dots, n \end{aligned} \tag{66}$$

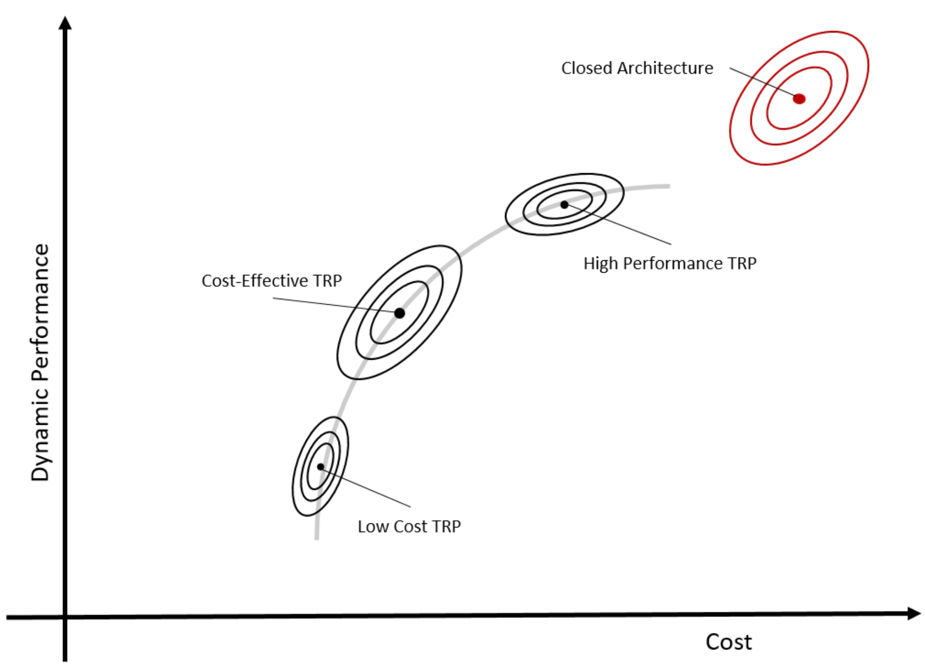


Figure 66: MCS Algorithm for Sensitivity Analyses with Flexibility and Uncertainty

The output of each iteration is the vector of dynamic performance and TLCC that would be realized if decision-makers followed an optimal course of action under the scenario created by the MCS. Merging these vectors across all trials then provides an empirical, joint CDF that is similar to the outcome of the JPDM process, albeit

www.manaraa.com

without the use of probability of success as an objective function. The remaining step is to fit a set of iso-probability contours to the empirical data that are consistent with the levels of interest to decision-makers. Plotting these iso-probability contours then provides an equivalent representation to the traditional cost-effectiveness plot characterized by Figure 9 in Chapter Two. In addition, it is possible to perform the same MCS on a closed system architecture concept, with the understanding that the closed system performance is tied to the terminal requirements and therefore remains unchanged in each iteration. If the results of the closed architecture MCS are provided on the same axes as the results from the open architecture analysis, then it is possible to directly compare the open and closed architectures under the existing cost-effectiveness framework. Figure 66 depicts a notional example of such an output for clarity.

Finally, the ability to evaluate alternative plans in the presence of both uncertainty and flexibility is the stated requirement for the sixth step of the proposed methodology. Figure 68 updates the process model with this observation.

المنارة للاستشارات
www.manaraa.com

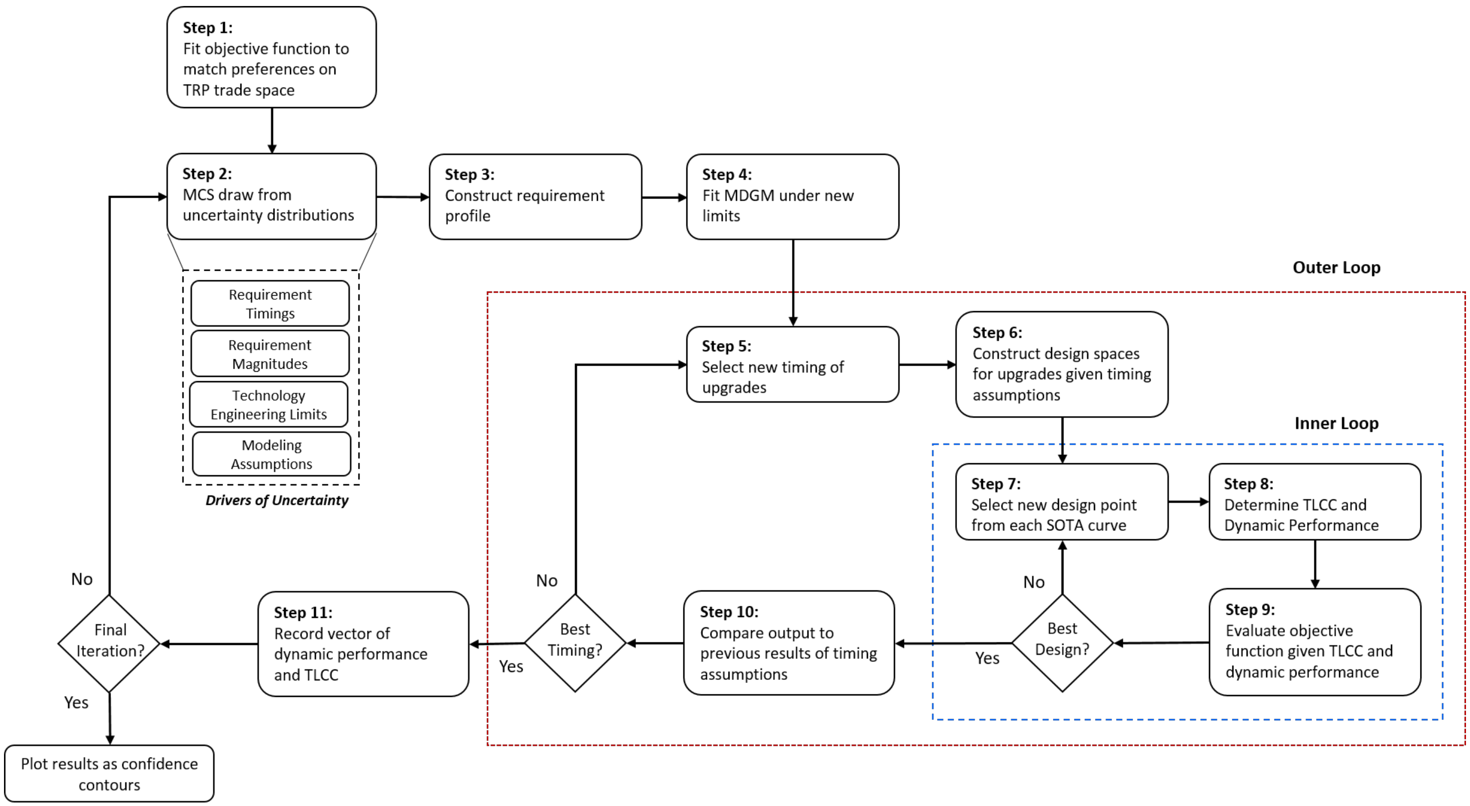


Figure 67: MCS Algorithm for Sensitivity Analyses with Flexibility and Uncertainty

المنارة للاستشارات
www.manaraa.com

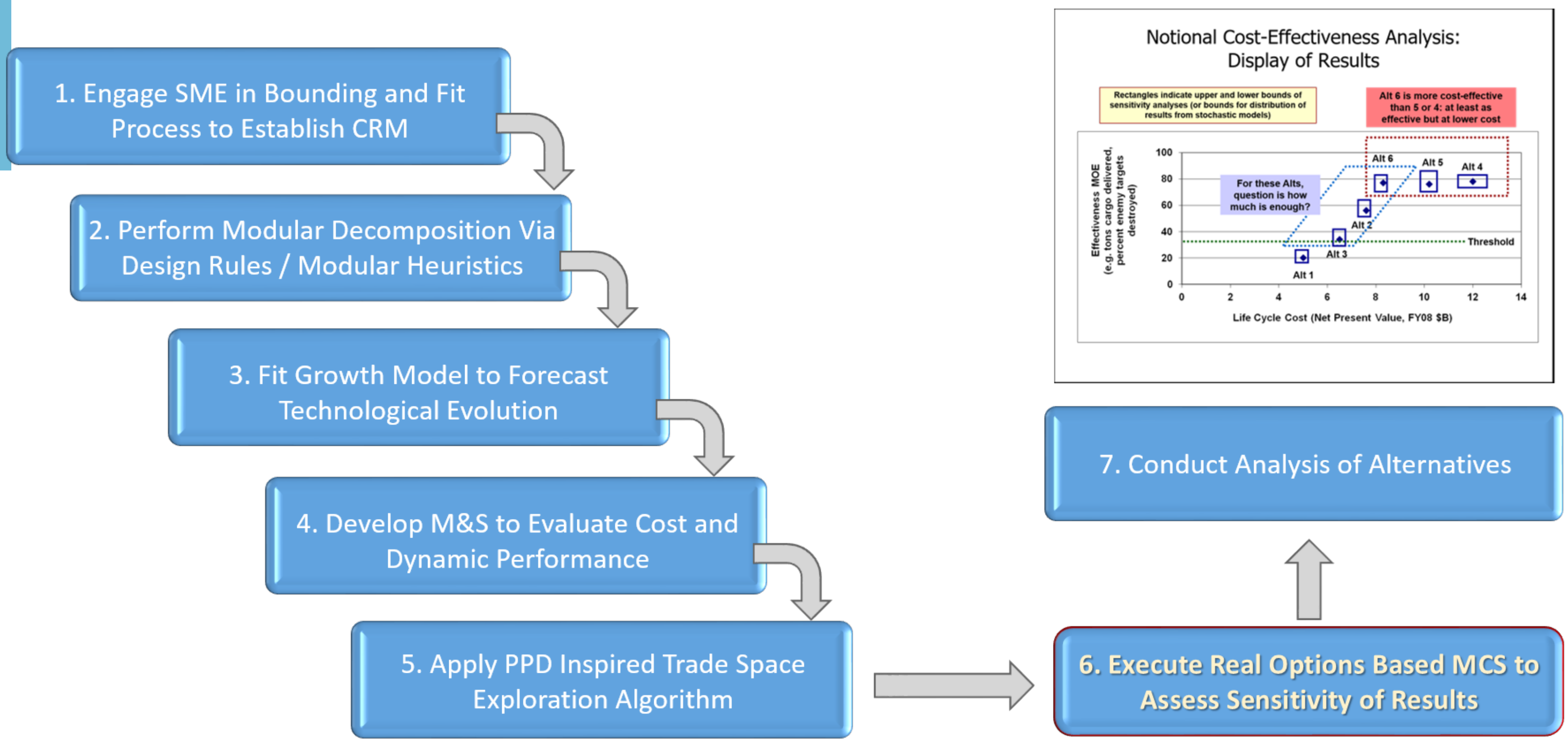


Figure 68: Methodology Update: Step Six - Real Options Inspired Framework to Evaluate Competing Alternatives

www.manaraa.com

CHAPTER VII

HIGH LEVEL SUMMARY AND OVERVIEW

With the Real Options heuristic framework for modeling the interacting forces of uncertainty and flexibility in place, it is now possible to fully describe the methodology. Recall that the methodology was developed as a response to the stated needs of the acquisitions community in resolving implementation barriers that prevent the conceptual advantages of OSA designs from translating into practical results. Chapter Two presented a review of SE methods developed by the acquisitions community to manage these challenges. Comparing these methods to the corresponding deficiencies observed by the stakeholder then provided insight into the gaps that must be addressed. Finally, these gaps were decomposed into three high level research questions that must necessarily be resolved to successfully satisfy the overall objective of greater OSA design implementation.

The first such question concerned bias mechanisms that, if not properly addressed, sow the seeds of failure in the earliest stages of design. Though there were many possible bias mechanisms, this work focused on those mechanisms that engineers can, and therefore should, address with rigorous SE methods: proper definition of system requirements, and development of a comprehensive modular partitioning scheme.

In terms of defining system requirements, it was determined that the existing "best practice" of documenting the evolution of requirements over time as a CRM is necessary, but not sufficient. This lack of sufficiency stems from the fact that there is substantial uncertainty inherent to these forecasts, and the existing methods of modeling this uncertainty were too qualitative and subjective for decision-makers to have confidence in the results. Chapter Four reviewed several quantitative approaches

المنارة للاستشارات
www.manaraa.com

found in the literature that could provide greater transparency and objectivity in this analysis. This review concluded that the best approach is to record uncertainty in requirement evolution as a set of CDF's defining the Functional SME's opinion on variability in the timing and magnitude of requirement increases. In addition, the process of soliciting this opinion and converting it into a mathematical form was codified in an iterative linguistic bounding and fit process. It is asserted that this process provides the requisite consistency, traceability, and objectivity to instill confidence in the results of this methodology.

With respect to modularity, it was observed that there are multiple methods to measure the extent to which modularity is present in a given partitioning scheme. However, there is no academic consensus as to how one would determine which method is the most appropriate, nor is there an accepted threshold to determine if a given score is acceptable. In other words, these methods allow the analyst to state that one partitioning scheme is more modular than another, but there are no established benchmarks to state that the more modular design has "enough" modularity to be deemed "modular". If one excludes quantitative metrics as a means of satisfying the research question, then the remaining schools of thought are the heuristic and DSM approaches for modular design. Both approaches have their relative advantages and drawbacks; heuristic methods ensure that all useful modules are identified, but scale poorly with increasing complexity; DSM methods manage complexity well by decomposing the system into isolated regions, but do not provide a consistent means of identifying modules within these isolated regions. The conclusion from this review was that an acceptable approach to modularity in an OSA design must be consistent (i.e. provide a repeatable process), scaleable (i.e. effort scales linearly with complexity), and complete (i.e. ensures all useful modules are identified). To that end, this work advocates a two-phase approach. The first phase applies Design Rules to partition the system at a high level, and the second phase applies heuristics to each isolated region

المنارة للاستشارات
www.manaraa.com

in order to identify a modular basis. This modular basis can then be evaluated against against the existing CRM's. If any of the modules provide a functional requirement documented in a CRM, then that module is set aside for further evaluation. Though this approach is asserted to meet the criteria, it should be noted that there is no "silver bullet" to specify the appropriate design rules and heuristics required to manage this process. Rather, these are the considerations where engineers should exercise their experience/expertise, and acquisition decision-makers should exercise their oversight authority.

The second research question began with the recognition that acquisition decision-makers will not accept that an OSA will be capable of efficiently satisfying evolving requirements on faith alone. Rather, open systems must undergo the same cost-effectiveness scrutiny to which closed system architectures are subjected. This requires, among other things, that a TRP be presented alongside the system concept to describe when components will be upgraded, what the resulting effectiveness will be relative to projected thresholds, and what TLCC should be expected given that these upgrades occur. Deductive reasoning led to the conclusion that developing a TRP to address these considerations first required the following to be resolved: how does one forecast the evolving properties of maturing technologies, and how can the complexities of time variation in the trade space of TRP's be reduced to a manageable representation.

Methods for predicting the evolving properties of maturing technologies come from the greater domain of Technology Forecasting. It was previous shown that forecasting methods can be categorized according the number of attributes and/or the dependency between attributes. As this methodology is intended to be system agnostic, those methods (i.e. univariate) restricted to the assumption of either a single performance attribute, or a set of mutually independent attributes were omitted from

www.manaraa.com

consideration. The remaining approaches were then evaluated in terms of their accuracy and the similarity of errors to a Normal distribution. The result of this analysis allowed the field of alternatives to be further refined to two approaches - TFDEA and MDGM. The MDGM approach was eventually chosen because it provided an efficient, closed-form solution for the SOTA curve at a future time, t, given by the formulation in Equation 35. This closed-form solution provides the first of two critical elements that allow for rapid exploration of the TRP trade space.

The second challenge, reducing the complexity of the trade space, proved to be more problematic. It was observed that the inclusion of the time domain as an additional degree of freedom excluded the possibility of leveraging the concept of Pareto efficiency to reduce the alternative space to a more concise trade space. Value Theory was considered as a potential path to resolve this conundrum, but ultimately proved to be inapplicable to the OSA design problem as there was still no apparent mechanism to consider time variation in the trade space. However, consideration of Value Theory eventually led to the identification of the related field of Dynamic Value Theory. Dynamic Value Theory is governed by two fundamental assertions: (1) the instantaneous value of a system is defined as the difference between what the stakeholders desire and what the system can deliver at a given point in time, and (2) stakeholders base their decisions on the life cycle value of the system, where life cycle value is defined as the integration of instantaneous value over the period of interest. The native form of Dynamic Value Theory, which is predicated on cost-benefit analysis, proved to be incompatible with the OSA design problem. Yet, if one recognizes an analogy where "what the stake holder desires" is contained in the requirement profile and "what the system can deliver" is contained in the performance profile, then the concept of Dynamic Value can be applied to the existing formulation. In so doing, it now becomes possible to reduce time variation in the TRP alternative space to an n-dimensional vector of scalar performance attributes. Similar to traditional Value

المنارة للاستشارات
www.manaraa.com

Theory, however, there are numerous alternative dynamic value formulations that can be leveraged. Chapter 5 presents several alternative models, but this represents another area in which the analyst can leverage their creativity. The only constraint is that the fundamental assertions of Dynamic Value be respected. Regardless of the specific formula applied, the Dynamic Value approach provides the second of two critical elements that allow for rapid exploration of the TRP trade space.

In review, a TRP is defined by a description of when the system is upgraded, what new technology is assumed to be infused, and the resulting TLCC. MDGM's provide the ability to determine what technology is believed to be available at a future point, and, given a set of assumptions on the timing/selection of technology, Dynamic Value provides a transformation to map the TRP into a vector of performance attributes. If this new perspective on the OSA design problem is accepted, then it becomes possible to evaluate and present the trade space of TRP's in a similar format to the existing cost-effectiveness framework. The singular distinction is that the performance axes reflect dynamic, as opposed to static, performance. The trade space exploration process, which was developed by applying lessons learned from the PPD literature, is presented in Figure 57. The framework uses a bi-level formulation, where the outer loop iterates through alternative timings, and the inner loop applies the NSGA-II algorithm to the design space. Each iteration produces a trade space corresponding to the specific timing assumptions generated by the outer loop. By filtering out inefficient points in between each iteration of the outer loop, the exploration process is believed to converge to a close approximation of the true dynamic frontier.

The element lacking in this trade space exploration process is proper consideration of uncertainty. The challenge, as noted by Research Question 3, is dealing with the interaction between uncertainty in future events and the flexibility afforded to decision-makers by an OSA to alter decisions over time. Chapter 6 presented a review of common methods currently in use in MODM literature to manage uncertainty.

www.manaraa.com

While these methods have broad appeal to static design problems, they are generally predicated on the assumption that uncertainty exists, and the best course of action is to find a design point that minimizes any sensitivity to this uncertainty. As such, there is no clear mechanism through which the concept flexibility as a hedge against uncertainty can be integrated. PPD design literature presents a class of methods that do explicitly consider flexibility in the decision-making process, but the event tree formulation does not appear to be applicable to the OSA design formulation. Given that the PPD methods for considering flexibility were drawn from Real Options, the literature on Options Theory seemed to the next logical source of inspiration. The results were mixed. On the one hand, there is a strong analogy between the attributes of an OSA design, having the right without the corresponding obligation to upgrade, and the type of design options considered by Real Options. On the other hand, Real Options methods are universally predicated on a cost-benefit system, which is incompatible with an acquisitions context for design.

The solution to this problem came through the recognition that many authors have begun recognizing a dichotomy within options literature. One way of thinking is to apply the specific the ROV techniques to problems that fit the context for which the valuation method was constructed. The other vein of thinking argues that when the specifics of ROV do match the problem at hand, such as the case with MODM problems, the appropriate course of action is to apply the real options heuristics that are embedded in ROV techniques. In considering the ROV techniques that, albeit tangentially, fit the OSA design context (i.e. multiple, discrete exercise opportunities), it became clear that the dominant heuristic is the pricing by simulation technique. This heuristic states that the proper course of action is to use an MCS to simulate alternative versions of future events that are consistent with present-day beliefs regarding uncertainty. Given that the evolution of future events is assumed to be known, one can then find the "best" course of action. As all scenarios are

المنارة للاستشارات
www.manaraa.com

equally likely, the distribution of results reflects the true uncertainty in results. The remaining question, then, is how one determines what the "best" course of action should be. To resolve this, it is assumed that the outcome of the deterministic trade space exploration process is a set of efficient points that decision-makers believe could be potentially useful. If these points are known, then a traditional objective function can be found with weights leading to the same conclusion. In other words, these objective functions simulate the decision-maker's preferences *a posteriori*.

The process for conducting this MCS is structured as follows. The MCS draws from each of the distributions modeling uncertainty in various aspects of technology evolution, requirement profiles, and modeling assumptions. Given this scenario, the outer loop assumes a set of timings for the infusion of technology, and passes these timings to the inner loop. The inner loops generates the MDGM corresponding to the assumed timings, and searches for the design points maximizing the decision-maker's *a posteriori* objective function. The result is returned to the outer loop, which perturbs the timing and repeats until converged. The vector of dynamic performance results and TLCC is recorded, and the next iteration of the MCS repeats. Once the number of MCS runs are exhausted, the results are tabulated as an empirical, joint CDF. This CDF can then be represented as a center of mass estimate and corresponding confidence bounds. This result is an identical formulation to the traditional-cost effectiveness plot presented in Figure 9, with the exception that performance axes again represent dynamic, as opposed to static, performance values. In addition, closed systems can be modeled under the same conditions, which allows open and closed architectures to be directly comparable under the existing decision support framework. This was the overarching objective of this methodology, and it is therefore asserted that the approach developed in this document meets this objective.

With this review in mind, the steps to apply this methodology are summarized in Figure 69.

المنارة للاستشارات
www.manaraa.com

The process begins with the formulation of a CRM. Developing this document will require the PM to engage the appropriate Functional SME's in the iterative linguistic bounding and fit process described in section 4.1. The end state of this process is a set of mutually agreed upon CDF's dictating uncertainty in the timing and magnitude of requirement increases. The next step in the methodology is to identify the specific system components that provide this capability. This is determined by construction of a DSM as a matrix representation of the lowest level of abstraction in a functional decomposition. Design Rules and heuristics are then applied in accordance with the guidance given in section 4.2 to identify the basis elements which compose the greater system architecture. Those elements contributing to the functional capability described in the CRM are then set aside for further consideration. The third step reviews the products available in the broader commercial market to determine if these physical components provide the functional attributes corresponding to the system elements identified in step two. If such a component exists, then a commercial product survey is undertaken. This survey identifies the relevant technology parameters by which the functional attribute can be measured, and collects data for the past progress of these parameters over time. Finally, a regression is performed to both fit the MDGM and to identify the projected engineering limits.

Once the growth models are fit, it will be necessary to establish an M&S framework to relate the technology parameters (DV) to the higher-level system objectives (RV) detailed in the CRM. This M&S must also provide a means to estimate the TLCC of the system, to include all relevant costs associated with generating, fielding, and supporting a component upgrade. Given that the future component is expected to provide an identical function with improved attributes, it is highly likely that a CER or a cost by analogy method will be the appropriate mechanism for estimating this cost. With the M&S complete, the analyst must identify an appropriate dynamic valuation formulation. This can be done by either selecting one the methods

www.manaraa.com

developed in this work, or by developing a problem-specific formulation. If a problem-specific formulation is developed, then the analyst must be careful to ensure that the assumptions used to satisfy the two fundamental assertions of Dynamic Value are well documented and approved by the appropriate decision-makers. At this point, it is now possible to identify the dynamic trade space using the methodology developed in Figure 57. Once established, the appropriate decision-makers must engage in an initial cost-effectiveness analysis to identify which design points warrant further consideration. A sensitivity analysis is then conducted for each of the points of interest. To facilitate this analysis, an objective function is constructed with a relative weighting scheme that matches any design points identified by decision-makers on the initial trade space. An MCS simulation is then performed to iterate through numerous possible outcomes in the evolution of requirements, technological progression, and modeling assumptions. At each iteration, the goal is to find the TRP that best matches the stated preferences of the decision-makers. These results are then tabulated, and converted into confidence bounds for each of the metrics in the dynamic trade space. The final step requires the decision-makers to reconvene for a second cost-effectiveness analysis in order to reevaluate the results in light of the confidence bounds. A final, informed decision can then be made regarding whether the open system approach is warranted and, if so, what TRP should be paired with the open design.

المنارة للاستشارات

www.manaraa.com

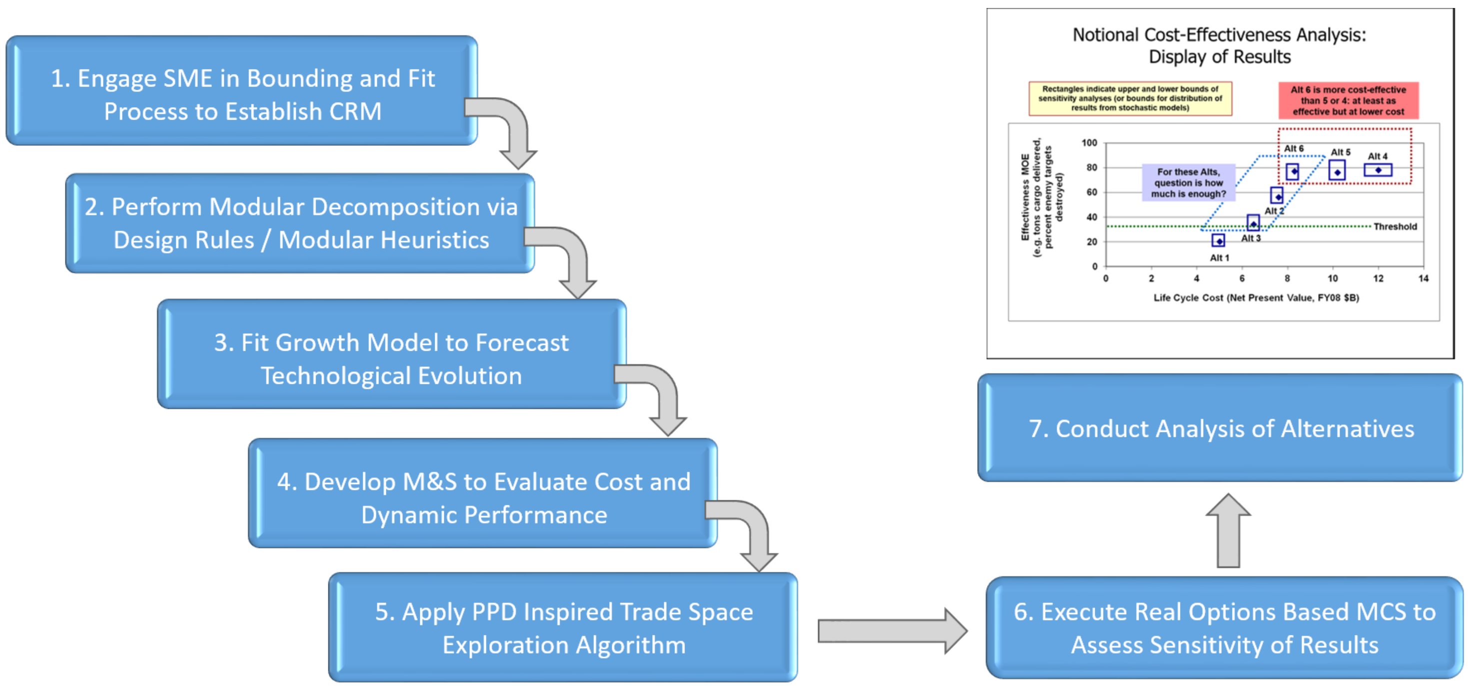


Figure 69: High-Level Overview of the Proposed Methodology

المنارة للاستشارات
www.manaraa.com

CHAPTER VIII

PROOF OF CONCEPT

Now that the proposed methodology is fully developed and described, the remaining task is to demonstrate that the methodology satisfies all of the aforementioned research objectives. This can best be accomplished by applying the proposed methodology to a real world problem where technological evolution provides the potential to efficiently satisfy evolving, uncertain requirements. With that in mind, this chapter will apply the methodology to a Intelligence, Surveillance, and Reconnaissance (ISR) design mission consistent with the original example taken from the Navy's Information Dominance Road Map. This objective statement is restated below [133]:

> Meet the growing demand coming from new Singnal Intelligence (SIGINT) and ocean-based sensors, as well as higher resolution persistent sensors (including FMV) coming from space-based systems and multi-spectral sensors...

The perspective taken on this scenario is that of the design team developing the aerial platform intended to host these improved sensors. Each incremental improvement in sensor technology will improve the ISR capabilities of the aircraft, but processing larger volumes of SIGINT data will also impose greater demands on the aircraft's Mission Computing Architecture (MCA). There are two architectural strategies that can be applied to manage these periodic increases in demand, which are presented below. The purpose of this analysis is not to determine which strategy is superior, but to assess the competing strategies in such a way that decision-makers can reasonably assess and compare their respective costs, benefits, and risks.

www.manaraa.com

1) Closed: Design sufficient excess capacity into the platform to provide high confidence of meeting all future demands

2) Open: Design to support near-term demand requirements, with the option to expand capacity later by infusing new technology

There are several reasons why this scenario is advantageous as an evaluation framework. First, the DoD Information Dominance team conducted a trade study on alternative combinations of technology that could be pursued to improve sensor performance. It will be shown that these results can be mapped to an equivalent processing requirement for the MCA, which is sufficient information to develop a real world CRM for the scenario. Next, the MCA's Central Processing Unit (CPU), the component directly responsible for satisfying computing requirements, is based on COTS technology with widely accepted measures of technical performance. Numerous products have been evaluated in terms of these performance metrics over the last decade, and the results of these benchmarking studies are available to the public. The availability of this data over a substantial period of time affords the ability to fit a forecasting model to real world, rather than notional, data. Finally, there are complimentary datasets for both the cost and SIGINT processing performance of various commercial CPU's over the same time period. The intersection of these three datasets enables the development of statistical regressions relating component attributes to cost and performance at the sub-system level. These regressions therefore serve as surrogate models for the CER and M&S requirements of the methodology. Thus, this scenario provides all of the necessary information to apply the methodology to a real world problem, with a minimal number of additional assumptions.

In conclusion, the proposed methodology was developed with a level of abstraction that makes it applicable to a wide variety of systems. If the methodology proves to be successful when applied to the real world scenario previously described, then that

www.manaraa.com

observation would lend considerable support to the argument that this approach is appropriate for the broader field of OSA design. Moreover, since this work was designed to be a sizable improvement over the current state of the art, support for the argument of wide scale applicability also supports the central thesis of this work - *the proposed methodology is a superior approach to address the shortcomings in modern OSA design.*

8.1 Step One - Establish a Capability Road Map

The Air Force Information Dominance team conducted a trade study in 2012 to determine how alternative combinations technologies could be leveraged to improved the performance of Wide Area Motion Imagery (WAMI) sensors [91]. These sensors are typically mounted on small Unmanned Ariel Vehicles (UAV) to provide a real-time bird's eye-view of the battle space. This provides an essential capability to initiate the F3EA mission cycle - Find, Fix, Finish, Exploit and Analyze - by identifying low signature targets, and relaying the actionable intelligence to ground forces. Figure 70 depicts a high level operational concept graphic for this mission set.

The trade study concludes with the identification of six "best value" alternatives that the authors argue are worthy of further development for future UAV capability improvements. From the platform designer's perspective, however, the specifics regarding these alternatives are irrelevant[1]. The question to be resolved is how much additional demand will these augmented payloads place on the processing capabilities of the MCA. From a computational standpoint, this demand is related to the number of pixels contained within the image (spatial resolution), and the rate at which the image is refreshed with new signals (temporal resolution). Each of the six alternatives identified by the authors was evaluated in terms of this aggregate demand, in units of Giga-Pixels per Second (GPS), and the results are provided in Figure 71 below.

[1]The interested reader is referred to the authors original work for a detailed discussion on the trade space process [91]

المنارة للاستشارات
www.manaraa.com

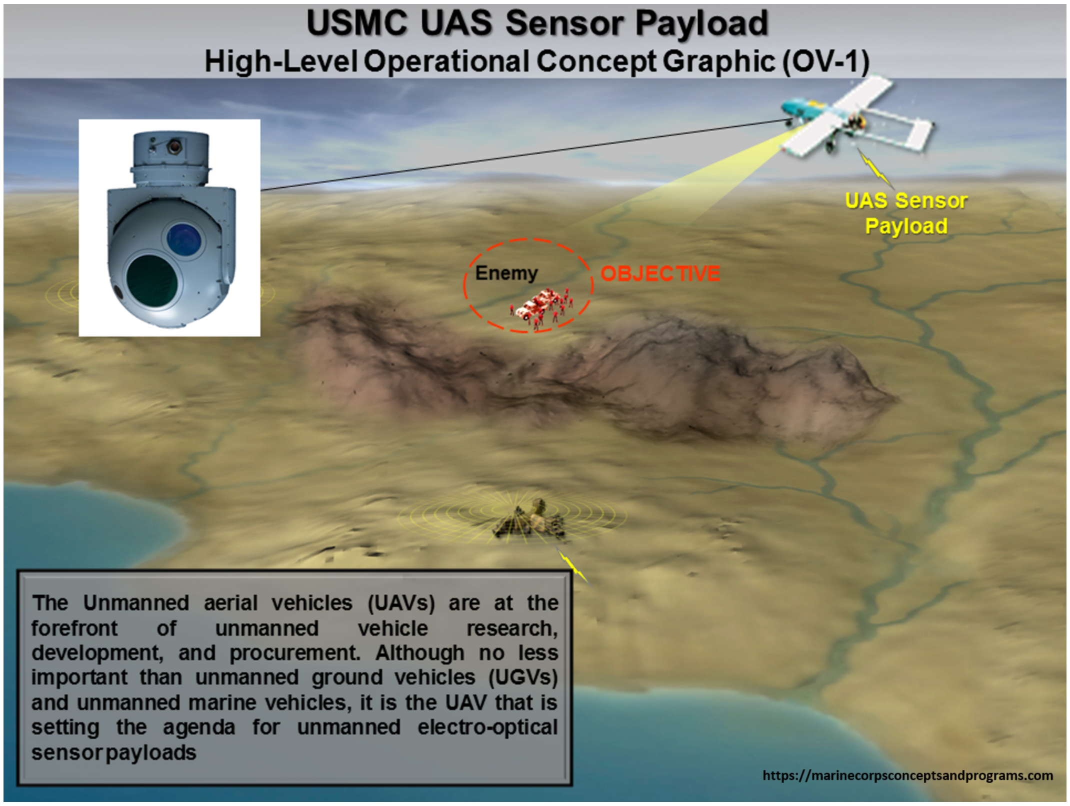


Figure 70: Operational Concept Graphic For UAS WAMI Reconnaissance

Figure 71 should be interpreted as follows. The blue vertical bars, which relate to the left axis, indicate the processing requirement that would be imposed if the corresponding alternative were developed and integrated into the aerial platform. Superimposed above these bars in red is the operational effectiveness score that, according to the author's value hierarchy, the corresponding alternative provides. These operational effectiveness scores are provided by the secondary axis on the right portion of the chart. With this in mind, two observations can be drawn from this analysis. First, higher levels of performance, and by extension greater operational effectiveness, impose larger computational burdens on the MCA. Second, Alternatives Two and Three require essentially the same processing bandwidth to accommodate, and the same is true for Alternatives Five and Six. Consequently, there are only four discrete settings that the platform could be called upon to provide in the event that one of the six sensor alternatives were to be developed and deployed.

المنارة للاستشارات
www.manaraa.com

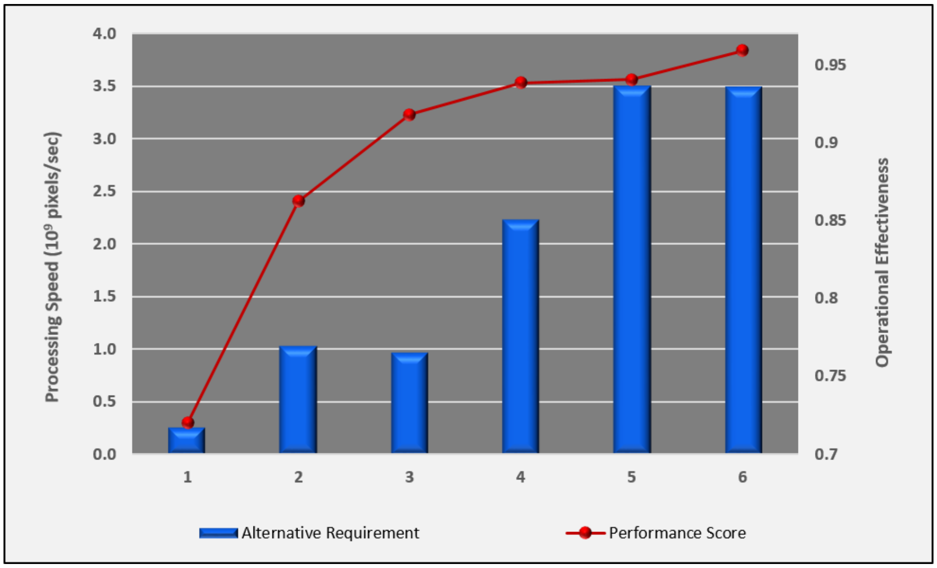


Figure 71: MCA Requirements Imposed By Augmented ISR Sensors

It is also important to note that greater levels of operational effectiveness can only be achieved through more aggressive combinations of advanced technologies. If one assumes that this implies greater commitments of time, resources, and development risk, then a realistic acquisition strategy would like follow an evolutionary approach. This approach would assert that those alternatives requiring less resources / risk be developed in parallel with the RDT&E required to construct the more complex alternatives. Under this paradigm, if the technological advancements required to produce a later variant prove to be too costly or difficult to achieve, then the DoD can terminate the program and still receive a useful capability increment. The scenario used in this experiment will assume that such a strategy is pursued, with the intent of migrating the system to progressively higher levels of operation effectiveness. These assumptions are summarized below, and depicted in Figure 72:

- System life cycle covers a 20 year period, from 2016 to 2036
- Initial capability of the system corresponds to Alternative One
- First capability increment will migrate the system to either Alternative Two or

المنارة للاستشارات
www.manaraa.com

Three in 2020

- Second capability increment will migrate the system to Alternative Four in 2025
- Third capability increment will migrate the system to either Alternative Five or Six in 2031
- Misc. software maintenance / upgrades increase requirements by 3% per year

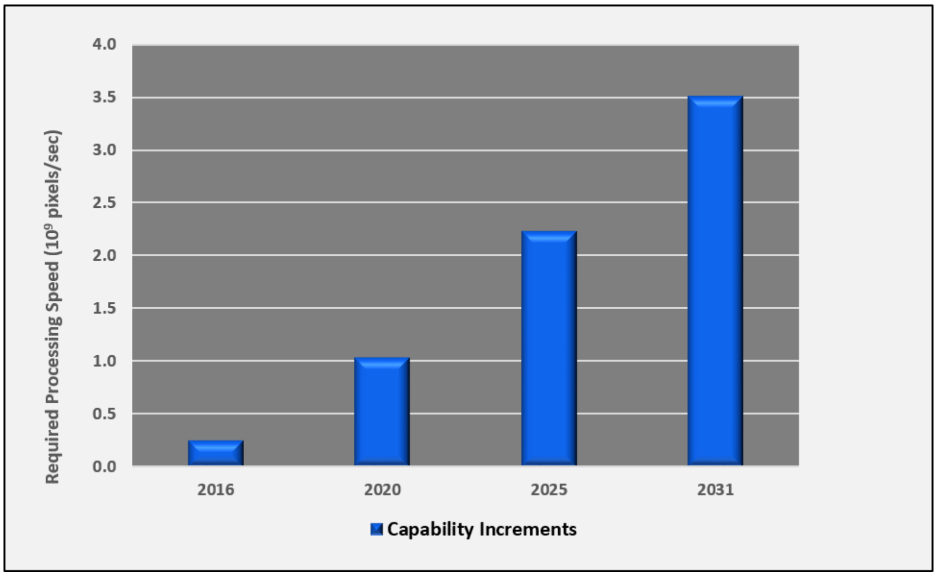


Figure 72: MCA Requirements Imposed By Augmented ISR Sensors

A challenge with the problem formulation up to this point is that the actual processing capabilities of the MCA are measured in terms of the number of Floating Point Operations Per Second (FLOPS) the processor can accommodate, whereas the present requirements are measured in terms of GPS. Establishing a relationship between the two quantities will require some consideration of the signal processing methods used to convert a continuous waveform to a binary artifact. One of the most common algorithms in this domain is the Fast Fourier Transform (FFT), which is simply a computationally efficient implementation of the more general Discrete Fourier Transform (DFT) algorithm. The purpose of this transformation is to convert a continuous signal from the time / spatial domain to the frequency domain, at which

www.manaraa.com

point filters can be applied to improve overall image quality and/or identify subtle indicators of concealed opponents. The principal benefit of the FFT approach is its ability to convert the complicated convolution operations required to implement these filters into a series of simple multiplications. This dramatically increases the speed with which filters can be applied to process data, but the algorithm is still quite time consuming to execute in real world applications [132, 155]. This scenario will therefore assume that the FFT transform is the limiting factor in the volume of date the MCA can accommodate at any given time. Under this assumption, there is a closed form solution for the number of FLOPS required to process an image consisting of N pixel values. This formulation is provided in Equation 67 [65].

$$FLOPS = 5Nlog_2(N) \tag{67}$$

Leveraging the transformation given by Equation 67 to convert the original requirement from GPS to FLOPS results in the MCA growth road map depicted in Figure 73, which provides a complete description of the timing and magnitude of requirement increases. As noted in Chapter Four, however, all such assumptions are inherently uncertain. Formulating the CRM therefore requires that each assumption be paired with a quantitative estimate of uncertainty. Chapter Four advocates that these measures are best ascertained by engaging Functional SME's in the linguistic bounding and fit procedure previous described. In the absence of a Functional SME, however, assumptions must be made to satisfy this requirement. To that end, it is further assumed that the uncertainty in each estimate is normally distributed with a mean corresponding to the transformed values provided by Figure 72. To determine the variance of these distributions, CDF bounds were constructed by assuming an upper and lower limit for each distribution. The variance of these distributions was then determined by applying the interior point optimization algorithm in Matlab to maximize the variance of each distribution, subject to the constraint that the

المنارة للاستشارات
www.manaraa.com

probability of selecting a value beyond the stated CDF bounds is less than 0.0001%. The bounds chosen for the various distributions and their associated parameters are provided in Table 12[2].

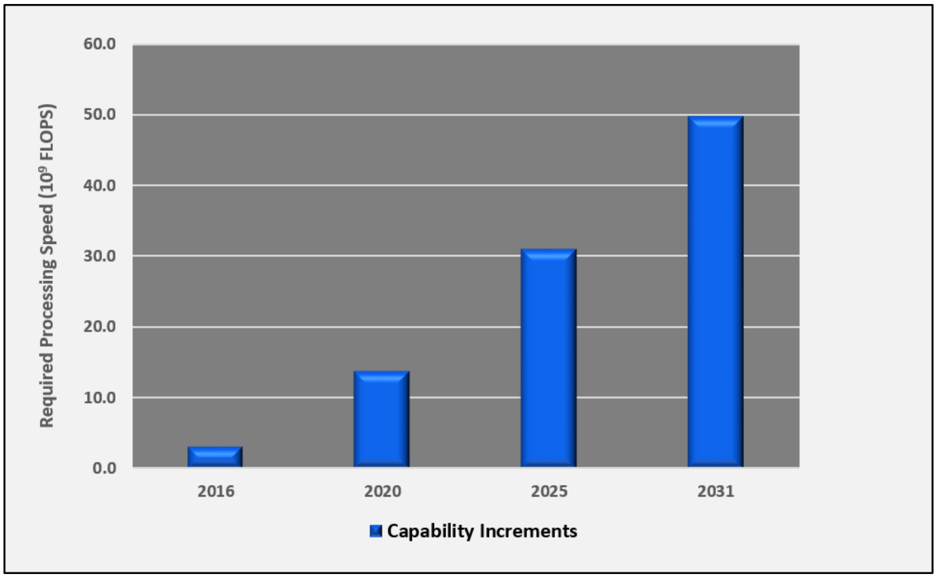


Figure 73: MCA Requirements Imposed By Augmented ISR Sensors

Table 12: Capability Road Map Parameters for Proof of Concept

	Requirements			**Timing**		
	Mean	Bounds	Std. Dev.	Mean	Bounds	Std. Dev.
Inc. One	10.66 GFLOPS	± 10%	0.3381	2020	± 0.5 years	0.1586
Inc. Two	17.31 GFLOPS	± 15%	0.8149	2025	± 1.5 years	0.4757
Inc. Three	18.69 GFLOPS	± 18%	1.0668	2031	± 3.0 years	0.9514
Annual	3.0%	± 1.0%	0.3171	NA	NA	NA

It should also be noted that the performance values provided in Table 12 represent the change in requirements from the previous increment, and the associated measures of uncertainty relate to this change. For example, if a MCS draw for the first requirement returned 11 GFLOPS, then the total system requirement at the next increment would be 11 GFLOPS plus the result of a MCS draw from the second

[2]The term *GFLOPS* in Table 12 is a convenient shorthand for Giga-FLOPS (10^9 FLOPS)

www.manaraa.com

performance distribution. This was a deliberate decision that was made due to the fact that the underlying technologies used to generate the three capability increases are not mutually independent. Therefore, if the first capability increment exceeded expectations, as in the example provided, then it is highly likely that future increments would also exceed expectations. The approach applied in this scenario loosely captures the propagation of these effects, though a more through treatment of this specific case study may lead to a different, and likely more sophisticated, model of this consideration. For clarity, Figure 74 provides a depiction of the CDF for the first performance increase that is consistent with the general form provided in Chapter Four. Graphs of the CDF's corresponding to the remaining assumptions are provided in Appendix C for further reference.

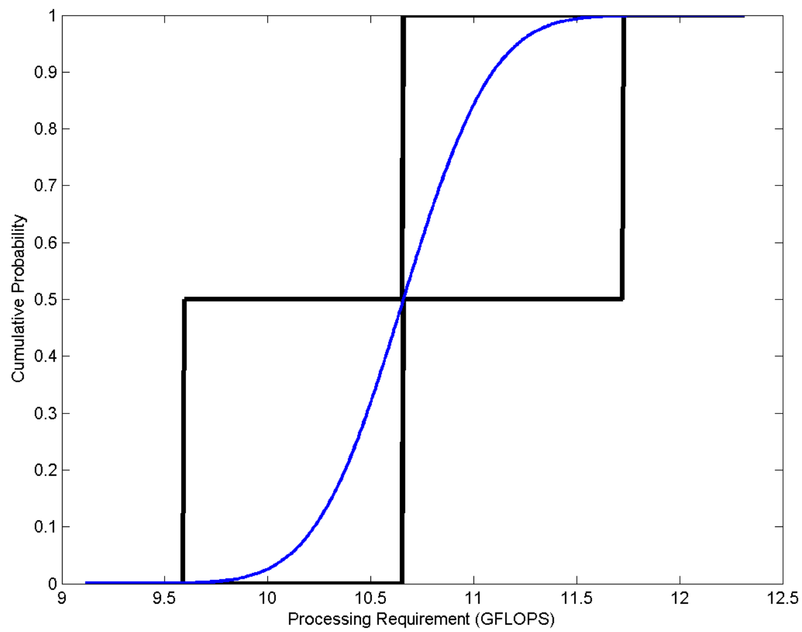


Figure 74: CDF of Requirements - First Capability Increment

With these measures of uncertainty in hand, it is now possible to express the CRM as the combination of an expected performance profile along its corresponding upper and lower confidence bounds. The bounds for this scenario were generated

www.manaraa.com

by performing 10,000 MCS iterations on the distributions identified in Table 12, the results of which are presented in Figure 75. The CRM generated by each simulation, the individual blue lines of Figure 75, was stored as a row entry in a matrix of running results, and the columns of the resulting matrix were sorted in ascending order. Finally, the 99% confidence bounds were extracted from the matrix by selecting the appropriate rows of the sorted matrix. The results from this second phase of analysis are provided in Figure 76.

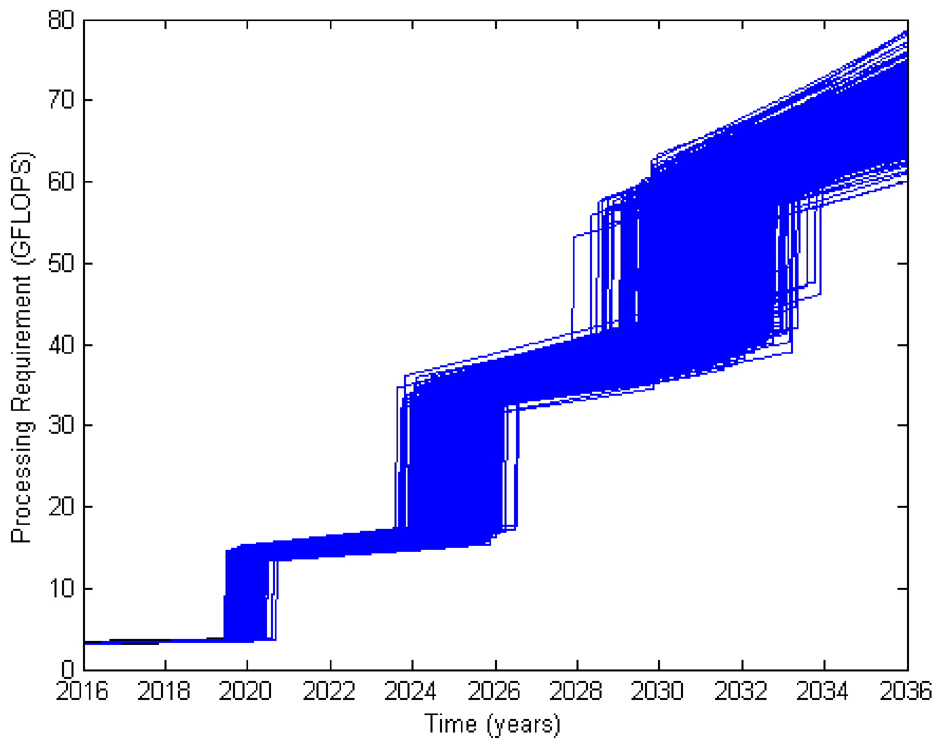


Figure 75: Results of MCS on Uncertainty Contained Within the CRM

Consider again the initial discussion surrounding the gold plating aspect of OSA design in Chapter One -*In order to ensure that the system is operationally useful throughout its planned service life, design requirements are derived from what is expected of the system at the end of its life cycle. Forecasting this far into the future, however, introduces significant uncertainty, and ensuring a high level of confidence that a design will satisfy its terminal requirements, given this high degree of*

www.manaraa.com

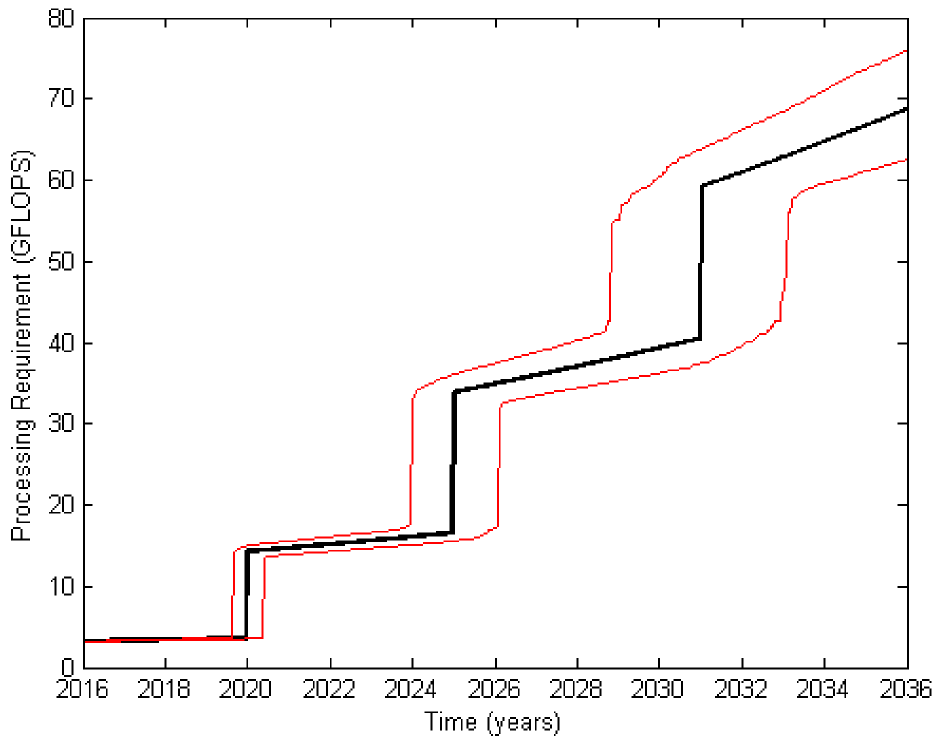


Figure 76: 99% Confidence Bounds of CRM

uncertainty, leads to the demanding requirements at the heart of the gold-plating phenomenon. With that in mind, Figure 76 clearly demonstrates that there is a cone of uncertainty surrounding the expected system requirements, where the diameter of the cone, and thus the magnitude of uncertainty, is monotonically increasing with time. Now imagine that the gold plating argument called for the upper 99% confidence bound to serve as the true requirement that must be satisfied in 2016. Under this scenario, the high confidence design would increase requirements by 10% relative to the expected profile, and 20% relative to the lower confidence bound. This is precisely the cause and effect relationship hypothesized by the gold plating argument. While this observation does not prove conclusively that the requirement generation process espoused by this methodology is correct, the fact that this concept is a byproduct of the modeling process does lend support to the central thesis of this work.

The final step in developing the CRM is to select the dynamic valuation concept

المنارة للاستشارات
www.manaraa.com

necessary to evaluate and compare competing TRP's. This scenario will apply the saturation template, with the assumption that an MCA cannot utilize processing capacity beyond 30% of its threshold requirement at any given time. In addition, instantaneous value will be scaled according to a utility template. The template assumes that satisfying the requirement threshold provides a minimum score of 0.5, while providing a level of capability at or beyond the saturation state provides a maximum score of 1.0. Performance values that lie between these two extremes will then be determined through a simple linear interpolation, which results in the utility function provided in Figure 77.

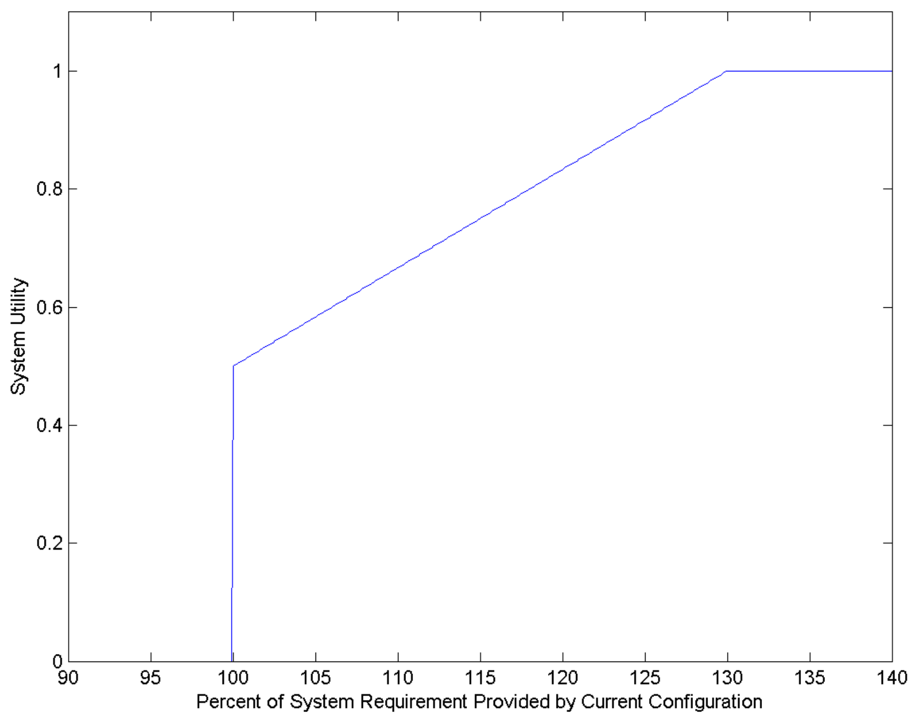


Figure 77: Utility Definition of Instantaneous Value

Under these assumptions, the CRM now consists of three properties: performance, system utility, and time. As such, the CRM can be modeled on three axes as a surface plot of alternative combinations in these properties. Figure 79 provides the surface plot corresponding to this scenario, along with a cross-sectional boundary imposed at

www.manaraa.com

a fixed performance value of 45 GFLOPS. The two-dimensional plot of time vs. system utility found at this cross-section is depicted in Figure 78, which presents an another interesting result: the utility of a system with static performance will steadily degrade over time until it reaches a point of technological obsolescence. This statement is an almost universally accepted truth within OSA design literature, but there are few, if any, mechanisms to quantify and model this property. Figure 78 clearly demonstrates that the proposed methodology provides a means to so, though this property was not deliberately considered when developing the requirement generation process. As with the previous observation, however, this result does not prove that the early stages of the methodology are *the* correct method to model the problem, but it does lend further credibility to the adequacy of the proposed approach.

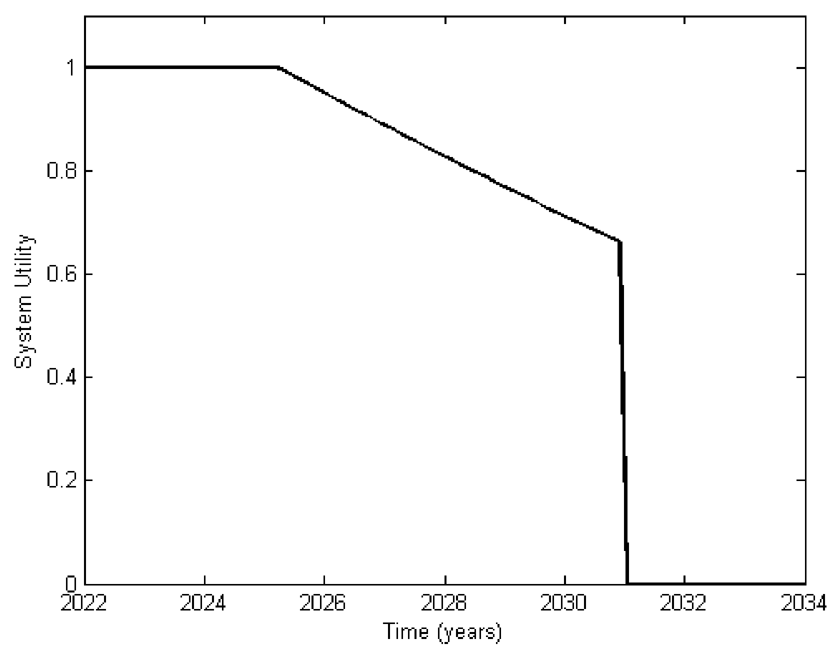


Figure 78: Cross-Section of CRM Surface Plot at a Fixed Performance Value

المنارة للاستشارات
www.manaraa.com

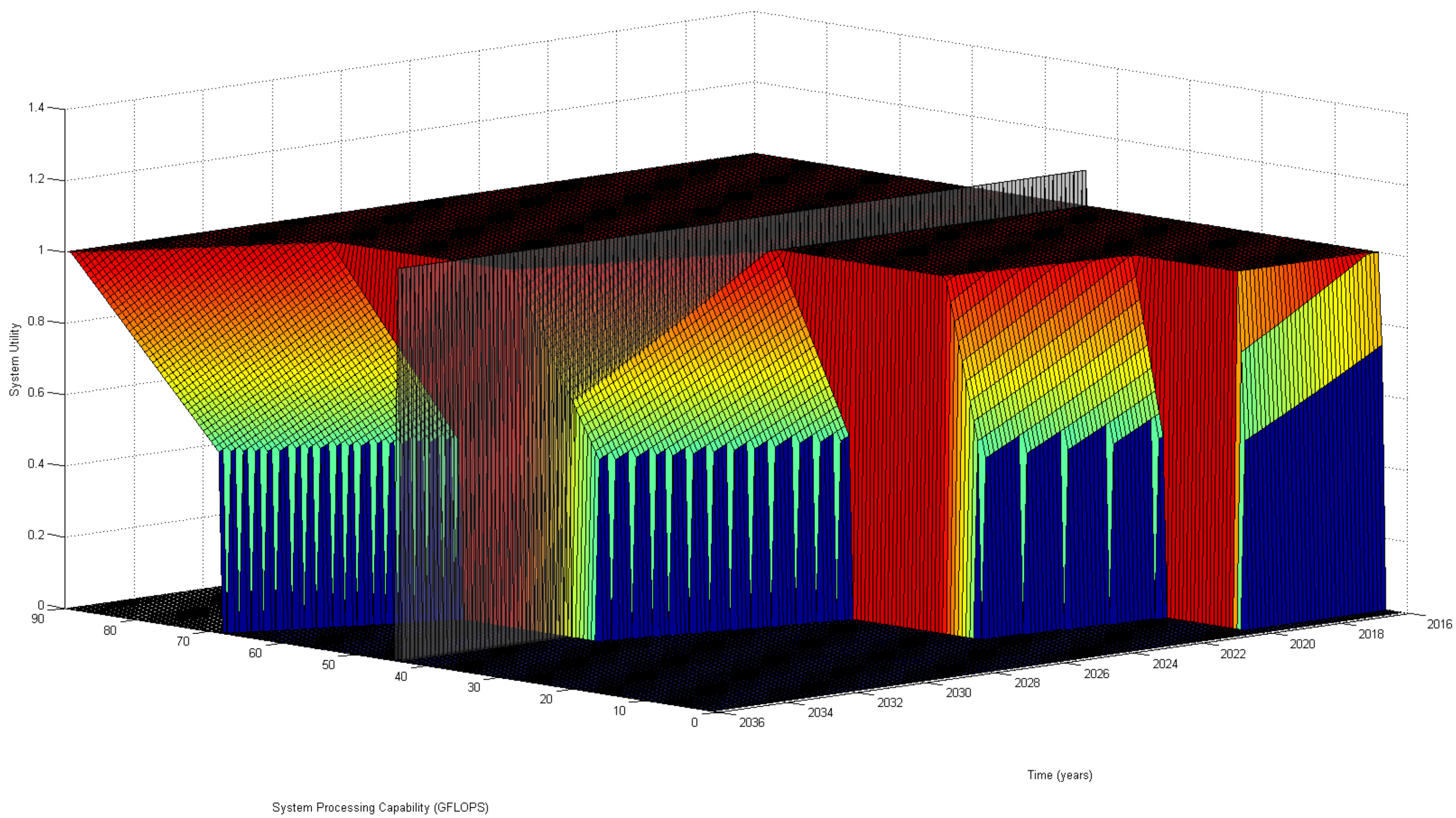


Figure 79: Surface Plot of the Utility Based CRM for the Experimental Scenario

www.manaraa.com

8.2 Step Two - System Decomposition

With the CRM fully defined, the next step in the methodology is to determine which portions of the system contribute to the higher level capability which is intended to grow over time. This work advocates a two phase approach as a baseline to manage this process, where the first step consists of developing Design Rules to partition the system into self-contained sub-systems. While this step is necessary to allow the methodology to be applicable to a generic system, it is not necessary for the experimental scenario described in the previous section since the desired capability is clearly contained within the MCA sub-system. For the sake of brevity, this scenario will therefore assume that the MCA is isolated from the greater system architecture to an extent that would prevent any component level changes from interacting with other major sub-systems.

The second step in this process is to then apply a set of relevant heuristics to further decompose major sub-systems into a basis set of modules. This type of decomposition was previously performed during a joint research effort between NAVAIR, the Georgia Institute of Technology, and the Georgia Tech. Research Institute to develop a COTS based variant of the existing, proprietary MCA design for the F-18. The sensitive nature of this work prevents a detailed description of this process, but it is sufficient to state that an MCA can be decomposed into a set of three major components[3].

- Single Board Computer (SBC): Complete computer built on a single circuit board. Contains a built in processor, memory, and data channels to handle a wide variety of general purpose tasks, to include signal processing.
- Graphics Processing Unit (GPU): Specialized component designed to rapidly manipulate and alter memory to accelerate the creation of images for output to

[3] Component descriptions adapted from https://en.wikipedia.org

www.manaraa.com

a display.

- Input/Output Board (I/O): Controls the flow of data between the SBC / GPU and the greater system architecture

Each of the three components connects to a backplane in a bus modular fashion, where the backplane serves as the interface connecting the MCA to the remainder of the system. Each of the three components listed above likely contributes to the overall processing capability requirements expressed in the CRM. This observation supports the necessity of including the second phase of modularization process, as a less rigorous approach may not identify all of the associated components. For the sake of simplicity, however, this scenario will focus exclusively on a single component, though it will be shown in a later section how simple modifications can be made to adapt the methodology for multiple components.

Curtiss-Wright VPX-187
Single Board Computer

Figure 80: Military Grade SBC in Used in Aircraft MCA Sub-Systems

The remaining question is which component should be considered. The most obvious choice would be to select the GPU, as its explicit function is to generate images from the FMV data stream. Yet, the assumption in this work is that the FFT

المنارة للاستشارات
www.manaraa.com

capability is used to process signals in order to identify subtle indicators of concealed components. This analytic functionality would likely be provided by the SBC, and it is therefore assumed that the SBC will serve as the sole upgradeable component for this scenario. Figure 80 provides an example of a COTS military grade SBC that will be the focal point for all subsequent analysis.

8.3 Step Three - Fit Growth Model

To fit a growth model to a given technology, one must first be able to idealize that technology as a vector of performance attributes. While this is a simple concept, it proved challenging to implement for SBC's. For example, the performance data sheet for the VPX-187 component illustrated in Figure 80 is 17 pages long and contains over 40 quantitative performance metrics. This baseline component also has a bewildering number of alternative configurations, each of which possesses some unique variation in the aforementioned attributes. In short, there is simply too much complexity in the native description of SBC components to reasonably accommodate the level of abstraction required for this methodology. A simplified representation is therefore required.

8.3.1 Parameter Identification

A simplified representation can be achieved by recognizing that the processing characteristics of the SBC, and thus their performance relative to the FFT capability requirement, are derived from their respective Central Processing Units (CPU). Within this category of sub-components, the 2006 performance benchmarks developed by the Standard Performance Evaluation Corporation (SPEC) are broadly accepted as the best comparative measure of performance across a wide range of hardware configurations. This is evidenced by the fact that over 8,000 individual SPEC benchmark tests were conducted in the last decade, most of which were performed by the original manufacturer. This benchmark suite uses four separate metrics to measure the

المنارة للاستشارات
www.manaraa.com

relative performance of commercial CPU's, and the results of these experiments are openly available at the SPEC website[4] [79]. The four metrics of interest are defined as follows:

- SPEC INT: Measures the time required for a computer to perform a single integer operation
- SPEC FP: Measures the time required for a computer to perform a single floating point operation
- SPEC INT Rate: Measures the number of integer calculations a computer can perform in a fixed amount of time
- SPEC FP Rate: Measures the number of floating point calculations a computer can perform in a fixed amount of time

There is, however, a competing argument in the literature for the use of a processor's theoretical FLOPS rate as an alternative metric for performance. This value is determined by multiplying the maximum number of parallel processes the CPU can support (i.e. operations per cycle) by the processor's clock-rate (i.e. cycles per second), and is therefore conceptually similar to the sizing variables common to other engineering analyses. To use a crude analogy, the maximum FLOPS rate can be viewed as the size of the pipe through which data flows, and the SPEC metrics can be viewed as a the speed of data transiting the pipe. Thus, if the size of the pipe is increased while holding the transition speed constant, then the overall processing performance of the CPU would increase. Both metrics should therefore be included in the analysis. Unfortunately, determining the exact value of this parameter for a given component proves to challenging, as it requires some degree of insight into the

[4]Test results and a more through explanation of the benchmarks are available at https://www.spec.org/cpu2006/

المنارة للاستشارات
www.manaraa.com

inner-workings of the processor's underlying architecture. The solution proposed for this experiment is to use the formulation provided by Equation 68. Here, T_i represents the theoretical FLOPS of the i^{th} component, c_i is the number of cores, t_i is the number of parallel threads per core, and ω_i is the clock speed. It should be noted that this formulation is not the true value, but it is proportional to the true value. This should be a sufficient approximation given that the performance parameter will be modified by a regression coefficient, which will scale the parameter to match response values. Still, this is an assumption, and its validity will need to reevaluated during the regression analysis in the next step of the methodology.

$$T_i = c_i \cdot t_i \cdot \omega_i \tag{68}$$

The discussion up to this point provides a means to idealize the performance of a CPU as a vector of five attributes: SPEC INT, SPEC FP, SPEC INT Rate, SPEC FP Rate, and T_i. In addition, a database of performance scores for real world products covering the period of time from 2006 to 2015 was identified. However, this database does not provide the corresponding date for the component's introduction into the commercial market, which is the response variable for the MDGM. A complimentary dataset was identified through CPU-World.com that contains the missing information, and the intersection of these datasets provides all of the necessary information to fit the MDGM[5].

8.3.2 Growth Model Fits

The general form of the Logistics MDGM is shown here as Equation 69, and is expanded in Equation 70 for the scenario currently under development. Danner, the original author of the MDGM approach, advocates that a useful step in this process is to linearize the general model under the transformation provided in Equation 72,

[5]It will still be necessary to identify cost and performances values for the components contained in this reduced dataset, but this will be addressed in the next section

المنارة للاستشارات
www.manaraa.com

which results in a final form given by Equation 71. This work will follow that example.

$$t = a - \sum_{i=1}^{n} \beta_i ln\left(\frac{L_i - x_i}{x_i - x_{o,i}}\right) \tag{69}$$

$$\begin{aligned} t = a &+ \beta_{theo} ln\left(\frac{L_{theo} - x_{theo}}{x_{theo} - x_{o,theo}}\right) + \beta_{int} ln\left(\frac{L_{int} - x_{int}}{x_{int} - x_{o,int}}\right) + \beta_{fp} ln\left(\frac{L_{fp} - x_{fp}}{x_{fp} - x_{o,fp}}\right) \\ &+ \beta_{intRate} ln\left(\frac{L_{intRate} - x_{intRate}}{x_{intRate} - x_{o,intRate}}\right) + \beta_{fpRate} ln\left(\frac{L_{fpRate} - x_{fpRate}}{x_{fpRate} - x_{o,fpRate}}\right) \end{aligned} \tag{70}$$

$$t = a - \sum_{i=1}^{n} \beta_i X_i \tag{71}$$

$$X_i = ln\left(\frac{L_i - x_i}{x_i - x_{o,i}}\right) \tag{72}$$

Applying the transformation in Equation 72 requires identifying the upper limit, L_i, and the starting point, $x_{o,i}$, for each performance parameter. In the case of L_i, the largest value found in each performance category is assumed to be 70% of the true engineering limit. For each $x_{o,i}$, a value slightly below the lowest value contained in the historical database was selected. In this way, the best and worst values found in the historical database serve, respectively, as the starting and limit values for the metric specific S-curves. Table 13 shows these limit and starting points alongside the best and worst values from which they are derived.

Table 13: Metric Bounds

	Starting	Limit	Worst in Database	Best in Database
Theoretic GFLOPS	2.0	105	3.6	73.6
SPEC INT	11.0	105	17.3	71.6
SPEC FP	13.0	168	19.2	112.5
SPEC INT Rate	21.0	1750	33.7	1231.5
SPEC FP Rate	17.0	1200	26.45	843.5

المنارة للاستشارات
www.manaraa.com

All design variables within the historical database were transformed according to the formulation provided by Equation 72 with the corresponding parameters given in Table 13. These transformed variables were then fit their respective response value according to the linear model defined by Equation 71. Before discussing the exact results of this fit, it will first be necessary to define the measures of "goodness" by which the results should be evaluated. Specifically, this research will focus on three primary considerations: the coefficient of multiple determination (R^2), residual plots, and parameter significance [45].

$\mathbf{R^2}$. The coefficient of multiple determination, R^2, is a mathematical measure to estimate how well the assumed functional form of the response measures variability in the supplied response data. This metric is ascertained through Equation 73, where SS_{Total} reflects the total sum squared error and SS_{Error} reflects the sum squared error that is directly attributable to the model. Clearly, a perfect fit of the data, one in which $SS_{Error} = 0$, would result in an R^2 value of one, whereas a complete lack of fit would lead to a value of zero. As a general rule, R^2 values for an acceptable model should be no less than 0.90, and preferably greater than 0.95. However, it should be noted that a high value of R^2 does not necessarily imply an acceptable fit, hence the need for the remaining two evaluation procedures, though a low value almost certainly indicates a poor fit.

In the context of this application, poor R^2 values can be an indicator of several problems. First, it is possible that some of the components identified in the historical dataset lagged behind the true state of the art, in which case the regression would attempt to fit points that do not lie on the true frontier. This scenario would be indicated by a set of points with a large, positive residual, which will be addressed shortly. In the event that this occurs, the resolution process would simply require the sub-standard components to be excluded, at which point the model should be refit and reevaluated against the reduced dataset. In the event that this trend is not

المنارة للاستشارات
www.manaraa.com

apparent, a low R^2 value may also indicate that an important performance metric was omitted from consideration. This would require a reevaluation of the metric identification process discussed in the previous section. Finally, if the model fails to provide a good fit after executing the previous troubleshooting steps, then the remaining conclusion is that the functional form of the MDGM is a poor representation of the underlying process. In this case, the assertion in Chapter Five that a MDGM is an acceptable method to forecast the state of the art in COTS components would be called into question. Moreover, since this assertion is a critical component of the proposed methodology, a failure of the MDGM would provide strong evidence that the method itself is sufficient.

$$R^2 = 1 - \frac{SS_{Error}}{SS_{Total}} \tag{73}$$

Residual Plots. The term "residual" in regression analysis is short hand for the difference between the actual regression values and those predicted by the model. This unexplained variation in the model is a form of error, and should therefore exhibit the same general principles of random error observed in real world environments: it should be normally distributed, with a mean of zero and a constant variance. Evaluating residual error in this way is best accomplished by reviewing a residual by predicted plot, where, as the name implies, the predicted value is presented on the abscissa and the magnitude of the residual is presented on the ordinate. An acceptable distribution of error would then be centered on a value of zero along the y-axis, indicating a mean of zero, with no observable trends in the scatter of data. The existence of an observable trend would likely indicate that higher order or non-linear effects exist, and are not captured in the functional form of the regression. In the event that this were true, is would again call into question the adequacy of the MDGM approach as an appropriate model of the state of the art.

Parameter significance. Each performance parameter within the MDGM should

www.manaraa.com

be assessed in terms of its importance in modeling the variability of responses. For a linear model, which is the focus of this work, the appropriate mechanisms to evaluate this significance are the t-statistic and p-value for the fit coefficients. For the purposes of this analysis, a coefficient is deemed to be statistically significant if the corresponding t-value is either greater than 2 or less than -2, and the p-value is less than 0.05. This ensures that there is at least a 97.5% confidence that the coefficient is non-zero. Any parameters that fail to meet this criteria are either not substantial indicators of the state of the art, or the historical database does not provide enough variation in the metric to adequately capture their significance [45]. Given the size of the baseline dataset, however, the former cause is far more likely than the latter. Consequently, any metrics failing to meet this criteria will be removed from further consideration.

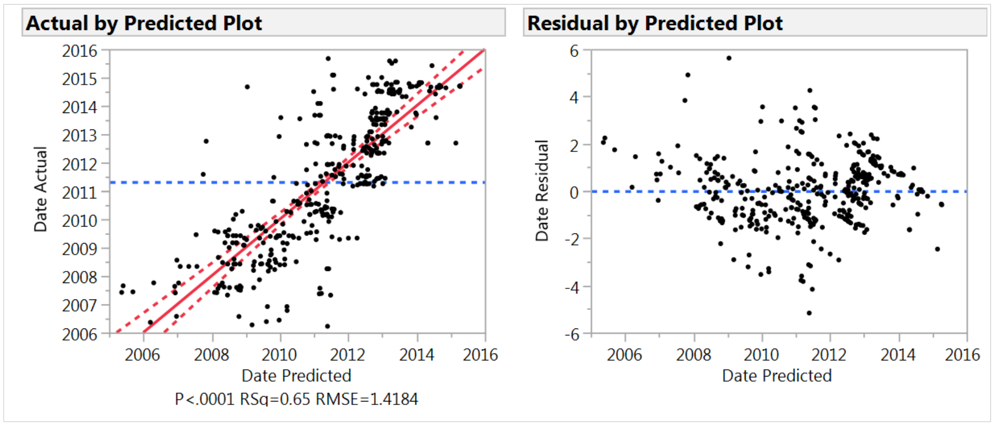
8.3.3 Evaluation of Results

The previous section established the metrics by which one should evaluate MDGM fits. The results of this fit for the entire commercial dataset are provided in Figure 81 below.

These results are, at best, dismal. The R^2 of this model is 0.65, which is well below the desirable threshold of 0.90. Considering the residual by predicted plot, there is a clear trend in the error structure where the residuals are stratified along straight lines. In addition, the extremes of the error distribution are on the order of 5 to 6 years, which is extreme given that the period of time covered by the dataset is only 10 years. Finally, the theoretical FLOPS and SPEC INT metrics fail to meet the standards of statistical significance. In short, the results of this model would seem to indicate that both the MDGM and the metrics identified in at the start of this section are entirely inappropriate to define the state of the art.

A closer review of the baseline dataset indicates a rather clear reason why the

المنارة للاستشارات
www.manaraa.com



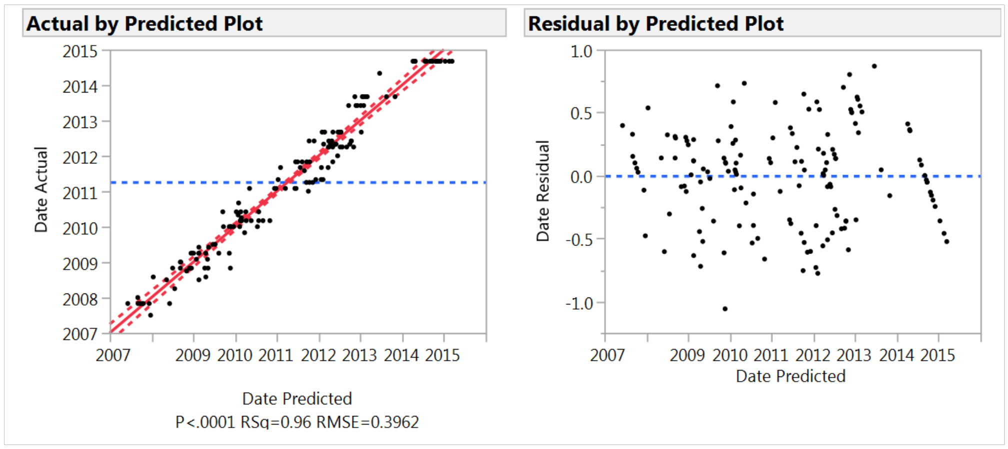
Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	2013.0715	0.199396	10096	<.0001*
FLOPS	-0.205802	0.126479	-1.63	0.1045
INT	-0.319244	0.314509	-1.02	0.3107
INTRate	3.6646282	0.461447	7.94	<.0001*
FPRate	-3.350702	0.462379	-7.25	<.0001*
FP	-1.295392	0.284421	-4.55	<.0001*

Figure 81: Results of MDGM Fit to Complete Dataset

initial model failed so spectacularly - it does not take into consideration cost. In other words, it is perfectly acceptable in the commercial market to release a product that pales in comparison to the current state of the art if it is available at a lower cost. To test this hypothesis, the baseline data was screened with a Pareto filter to remove those entries where a component could be found either at the same time or earlier with superior scores in all dimensions of performance. In total, this filter reduced the total size of the dataset from 379 to 144 distinct entries. The results of this reduced dataset are provided in Figure 82.

These results are significantly better. The R^2 value has increased from 0.65 to 0.96, which falls well within the ideal range to justify the functional form given by the MDGM. Moreover, the error structure has become much more random in its scatter across the residual by predicted plot, though there is still some degree of stratification that warrants further consideration. With respect to the individual performance measures, the theoretical FLOPS metric is now statistically significant, which supports the original justification for its inclusion within the model. The SPEC

المنارة للاستشارات
www.manaraa.com



Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	2017.6245	0.343084	5880.8	<.0001*
FLOPS	0.5069526	0.06265	8.09	<.0001*
INT	-0.151453	0.201356	-0.75	0.4532
FP	-2.485556	0.222814	-11.16	<.0001*
INTRate	2.8078138	0.321472	8.73	<.0001*
FPRate	-3.405707	0.354868	-9.60	<.0001*

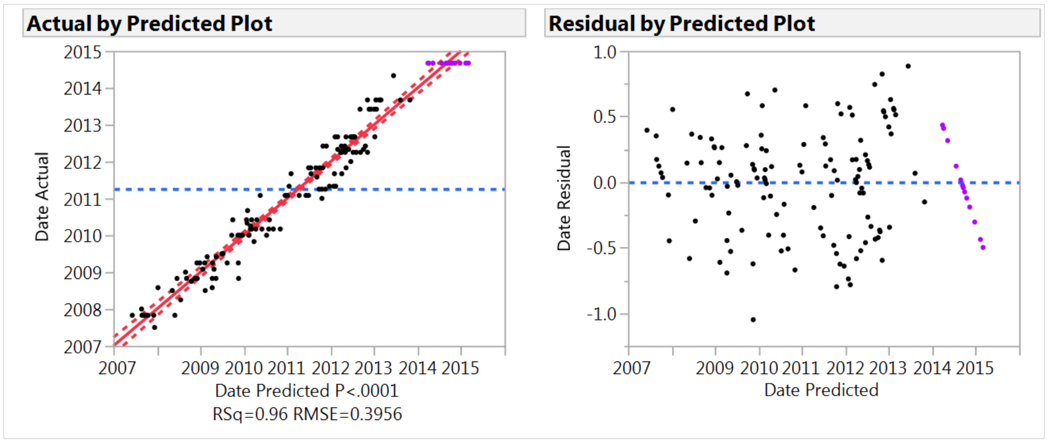
Figure 82: Results of MDGM Fit to Reduced Dataset

INT metric, however, still appears to be statistically insignificant, and can therefore be assumed to be irrelevant in determining the state of the art.

With this in mind, the MDGM was fit for a third time with the SPEC INT metric removed, and the results are depicted in Figure 83. In this final revision, the R^2 remains unchanged at 0.96, and all attributes are statistically significant. In addition, the highlighted points in both graphs, which represent the same entries, helps to explain the subtle stratification trend in the error structure. These points all correspond to the Intel Xeon-E5 component line, which were released on the same date as a product family. If one assumes that the commercial advantages of releasing products as a family supersedes the benefit of releasing individual components, then the impact on the timing of releases would explain the stratification of residuals.

In conclusion, the final MDGM formulation in Figure 83 satisfies all criteria for a successful implementation of the MDGM. This not only validates the metric identification process at the start of this section, it also provides substantial support for the technology forecasting arguments made in Chapter Five. Finally, by providing

www.manaraa.com



Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	2017.86	0.140138	14399	<.0001*
FLOPS	0.5172782	0.061032	8.48	<.0001*
FP	-2.647158	0.058949	-44.91	<.0001*
INTRate	2.6801497	0.272587	9.83	<.0001*
FPRate	-3.24834	0.286186	-11.35	<.0001*

Figure 83: Results of MDGM Fit to Reduced Dataset

support for the third step in the methodology, these results also contribute to the thesis that the proposed methodology is a superior approach for OSA design.

8.4 Step Four - Develop Modeling and Simulation

The purpose of M&S in this work is to map design variables to their associated responses. For the specific experimental scenario currently under development, the design variables are defined by the metrics identified in the previous section, and the responses of interest are unit cost and FFT performance. This section will develop these models and, in so doing, test more the assertion made in the development of the proposed methodology.

8.4.1 Cost Model

Chapter Five asserted that future components should be idealized as an identical copy of existing products, with the sole distinction that the performance attributes of the future component show some measure of improvement over their predecessor.

المنارة للاستشارات
www.manaraa.com

This implies that existing CER practices, where cost is estimated based on parameterized performance values, would be an acceptable means of assessing the cost of future components. The first step in developing a generic CER is to again compile an historical database, though in this instance all entries must reflect the cost of components at the date of their introduction into the market place. These entries must then be paired with their corresponding entries in the performance database in order to relate a specific vector of attributes to a specific cost. The remaining step is to then fit a regression relating design and response variables.

To develop the CER for this scenario, recall that the initial dataset used to fit the MDGM was developed by identifying the intersection between the SPEC dataset and the complimentary dataset from CPU-World.com. This was necessary because the SPEC dataset did not contain the necessary response variable, which, in that case, was the introduction date. The same argument exists with respect to cost: the SPEC dataset does not contain cost data, but the complimentary dataset, rather fortuitously, does. This allows the reduced dataset from the previous section to be reused for the present analysis. In addition, the initial failure of the MDGM fitting process was attributable to the large number of inefficient points contained within the historical database. Given that the same data will be used to fit the cost model, there is a high likelihood that the same observation would hold true. To test this hypothesis, the data set was screened with a second Pareto filter to remove those entries where a component could be found at an earlier date that possessed better performance in all categories and was available at a lower cost. In total, this filter reduced the total size of the dataset from 379 to 154 distinct entries, which is only slightly larger than performance efficient set identified in Step Three. This is an interesting result, as the hypothesis in MDGM section stated that cost was the dominate explanation for why there were so many performance inefficient products in the commercial market. In reality, however, this logic only represents 5% of the inefficient data, and it is

المنارة للاستشارات
www.manaraa.com

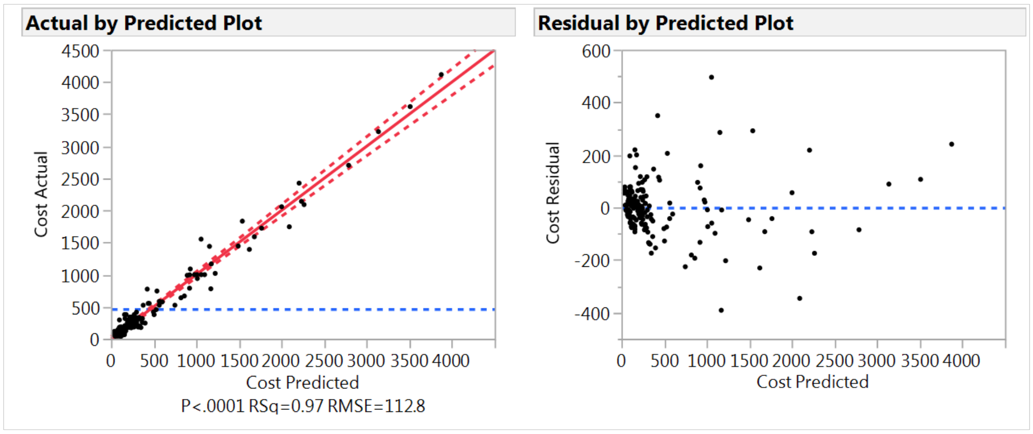
not clear what factors explain the remaining 95%. A final hypothesis that could explain this difference is the fact that commercial markets are extremely competitive, and manufacturers must develop products without perfect knowledge of the products being developed by their competitors. Evaluating this hypothesis, however, would require research that is beyond the scope of this work, and is therefore left as item of consideration for future work.

There is clearly a great deal of overlap between the development of the MDGM and CER, but a substantial point of departure is the general form of the regression to be fit. For the MDGM, the functional form was created to capture the hypothesized process of technological evolution described in Chapter Five. The CER, on the other hand, has no predefined form; this must be determined by the analyst. A reasonable start point when attempting to determine an appropriate of the CER is to begin with the generic Response Surface Methodology (RSM) template given by Equation 74. Here, β terms are regression coefficients, x_i represent component performance metrics, and ϵ is the model error. Once this initial model is fit, it can be further refined by removing any parameters failing to satisfy the statistical tests of significance. In addition, discernible patterns in the residual structure can imply that transformations of response variables (e.g. logarithmic transformation of responses prior to fitting the model), or higher order terms can be leveraged to further improve results.

$$cost = \beta_o + \sum_{i=1}^{m} \beta_i x_i + \sum_{i=1}^{m} \beta_{ii} x_i^2 + \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} \beta_{ij} x_i x_j + \epsilon \tag{74}$$

In executing this process, time was initially added to the existing performance metrics as a fifth design variable. This creates 21 fit terms in the baseline model, several of which proved to be statistically significant. Those metrics which account for the least variation in response were then sequentially removed from the model until all remaining terms satisfied the test of statistical significance. This reduction left 10 remaining terms, and the results are provided in Figure 84.

المنارة للاستشارات
www.manaraa.com



Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	220059.21	26594.01	8.27	<.0001
Date	-109.5184	13.2281	-8.28	<.0001
FP	6.2588	0.8357	7.49	<.0001
INTRate	3.9968	0.3794	10.53	<.0001
FPRate	-3.7999	0.4887	-7.78	<.0001
Date*Date	15.0193	4.5369	3.31	0.0012
FLOPS*FLOPS	0.8923	0.0671	13.30	<.0001
FLOPS*FP	-0.5931	0.0465	-12.75	<.0001
Date*IntRate	-0.5054	0.0732	-6.90	<.0001
FLOPS*INTRate	-0.0220	0.0040	-5.55	<.0001
FP*INTRate	0.0517	0.0037	14.08	<.0001

Figure 84: CER Fit with Time as Dependent Variable

The resulting model has an R^2 value of 0.97, and all parameters exceed the tests of statistical significance by a wide margin. There is an apparent clumping of points around the lower portion of the residual plot, which is a common indicator of underlying shortcomings in the proposed model. In this instance, however, that trend is to be expected, as there is an extremely uneven distribution of commercial products across the cost dimension of the historical database. In other words, low cost components are far more common in the commercial market place than their high performance, high cost counterparts. With that in mind, there is another potential pitfall that takes greater priority. This pitfall stems from the fact that there are several non-linear basis functions in the proposed model, and there is significant possibility that the quadratic terms in the model could become dominate when extrapolated forward in both time and performance. An extrapolation test should therefore be conducted to evaluate the significance of this effect.

To perform the extrapolation test, the model is evaluate under varying percentiles

المنارة للاستشارات
www.manaraa.com

of the performance attribute's upper limit at a discrete set of times. Figure 85 below depicts the results of this analysis from 10-100% of the upper performance limits in the years 2016, 2024, and 2032. If the model performed as desired, then the performance results should demonstrate a clear, monotonically increasing trend across the performance spectrum, and costs at each point should reduce as time increases. Clearly, this is not the case. The initial trend for 2016 (blue line) appear promising across the performance domain, but the trends in 2024 and 2032 are quite puzzling. In 2024, the cost of components decreases as performance scores are raised from 10% to 50% of their respective limit states and actually become negative, at which point the trend resumes a trajectory back into positive territory. In 2036, the situation is even worse, as the maximum cost is found at a minimum performance and the majority of prices are negative. Thus, even though the proposed model satisfied all regression tests for suitability, it is clearly not capable of extrapolation.

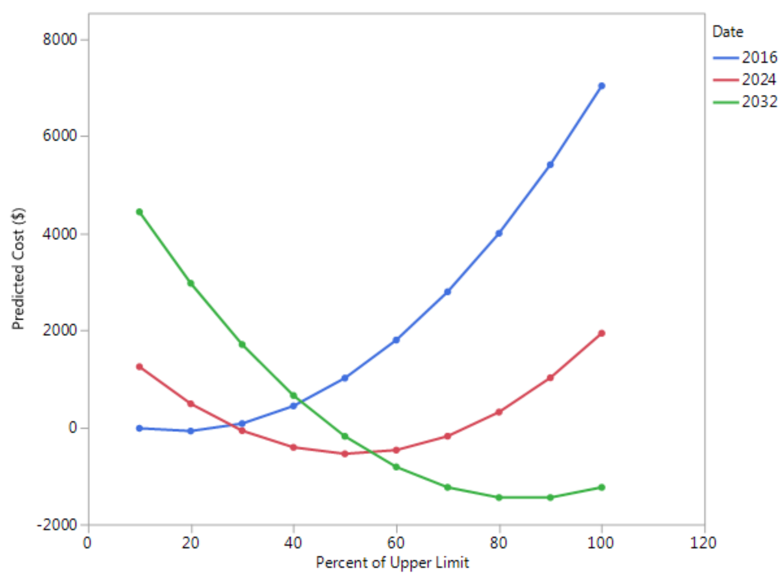


Figure 85: Extrapolation Study for the Initial CER - Performance and Time

In reviewing the parameters in Figure 84, this result should have been expected given that the coefficient for the quadratic time parameter is substantially larger than most other terms. Several alternative model configurations were attempted to

www.manaraa.com

resolve this dilemma, but no models were found to pass the extrapolation test where time is directly included as a dependent variable. Consequently, it would seem more appropriate to develop a model in which the impact of time is accounted for indirectly. To consider how such a model would be constructed, it is important to note that the prevailing assumption in OSA design is that as time progress, costs for an equivalent product will decline. If this assertion is true, then there should be historical evidence of this trend in the form of price inflation, or in this case deflation, over time that would allow the long term trend to be quantifiable. This data does in fact exist for computer hardware, and the long term trend is provided in Figure 86 below[6].

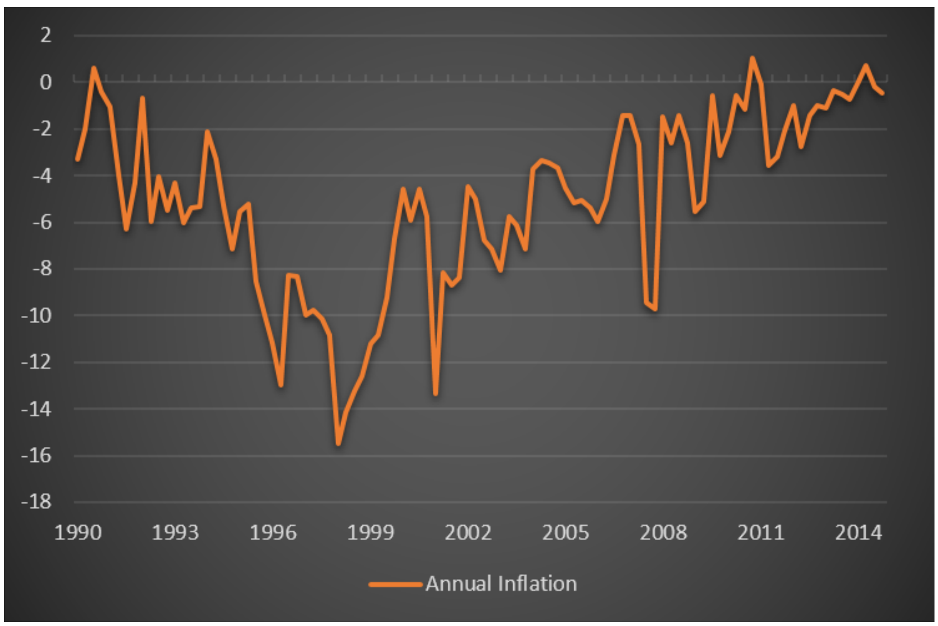


Figure 86: Anuual Price Deflation of Commercial Components

Based on the data presented in Figure 86, it will be assumed that prices decline by 2% per year, though this is an area of uncertainty that must be addressed in the sensitivity analysis. Applying this trend will first require a CER that uses strictly performance data to determine what the cost of a given component would be if it existed in the present, even if the associated performance values exceed the state of

[6]Data retrieved from U.S. Federal Reserve Survey: http://www.federalreserve.gov

المنارة للاستشارات
www.manaraa.com

the art. In other words, this is a determination of what the component would cost if it were available today. Once this is known, the present cost is deflated to the correct date using the standard discounting formulation provided in Equation 75. Here, F_t is the future cost of the component, PV is the present value (i.e. the cost in 2016) of the component, i is the deflation rate, and t is the future time at which the component would be available under the MDGM constraint on the state of the art.

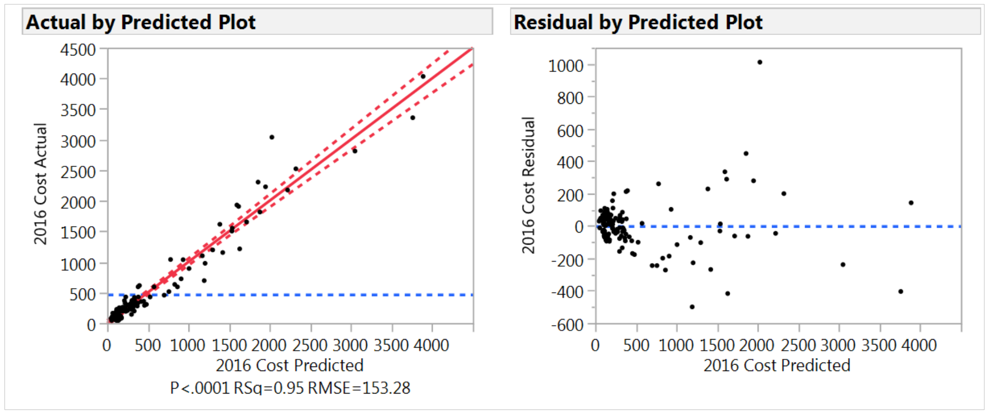
$$F_t = PV \cdot (1 - i)^{t-2016} \tag{75}$$

The next step is to develop the performance based CER necessary to determine a component's present value. This requires a slight modification to the existing data set, as all cost data must be inflated to a 2016 equivalent[7]. However, once this is achieved, the model development process proceeds as previously described. In this case, the initial RSM is composed of 15 basis functions, which were reduced to 6 after the parameter reduction procedure was implemented. The results of this fit are provided in Figure 87.

Given the challenges with the initial model, it would be prudent to conduct a similar extrapolation study on the new cost model. Results from this analysis are depicted in Figure 88, though in this instance there is only one curve since the extrapolation is contained within the performance dimension. A cursory inspection indicates that the results have improved substantially, as the cost model now demonstrates the monotonic property in which increasing performance increases costs. In addition, the limiting case in which component performances reach 100% of their respective limit states yields a cost of $7,361, which is approximately 50% greater than the maximum cost found in the historic database. This appears to be a reasonable result, which leads to the conclusion that the proposed model is sufficient to proceed.

[7]This was done by assuming a 2% inflation rate over the period of time covered by the historical data

المنارة للاستشارات
www.manaraa.com



Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-79.9986	18.009	-4.44	<.0001*
INTRate	1.8160	0.0787	23.07	<.0001
FLOPS*FLOPS	0.6481	0.0835	7.76	<.0001
FLOPS*INTRate	0.3459	0.0387	8.93	<.0001
FP*INTRate	-0.1257	0.0208	-6.03	<.0001
FLOPS*FPRate	-0.5135	0.0557	-9.21	<.0001
FP*FPRate	0.1800	0.0277	6.50	<.0001

Figure 87: CER Fit to Performance Metrics

8.4.2 FFT Performance Model

The second purpose of M&S in this formulation is to map the component design variables identified in the previous section to the performance measure contained within the CRM. Such a mapping is typically established through a physics based simulation, which may be replaced at a later stage with a surrogate regression model to expedite the design space exploration process. Unfortunately, there is no apparent physics based relationship between the design and response variables: SPEC performance metrics are relative, unit-less measures of performance, and there is no apparent way to determine the measure of efficiency that could relate a CPU's theoretical potential to its actual FFT performance. To circumvent this deficiency, another historical database was assembled to aggregate the results of various benchmarking tests used to directly evaluate the FFT performance of CPU's. Components represented in both databases, the new FFT performance database and the previous SPEC database, therefore provide a direct relationship between design and response variables. These observations can then used to develop the requisite surrogate model

www.manaraa.com

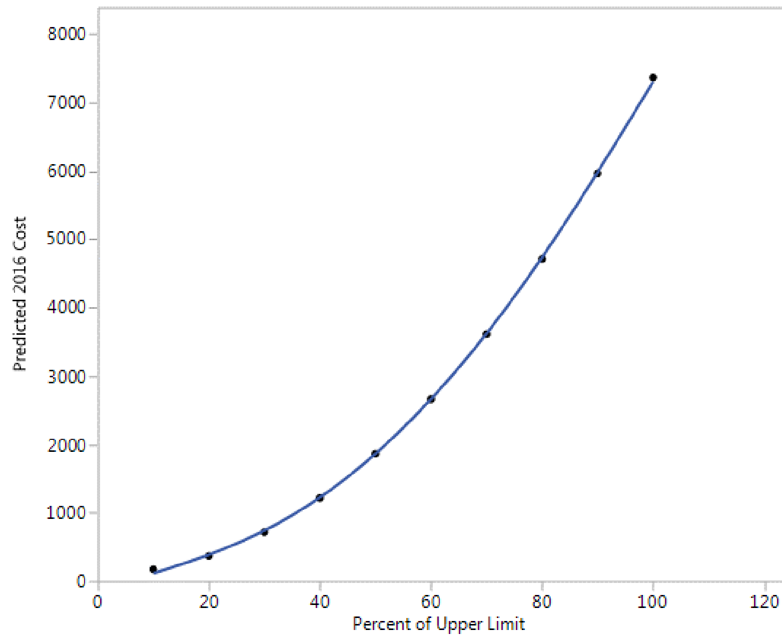


Figure 88: Extrapolation Study for the Revised CER - Performance Only

for rapid exploration, which essentially replaces the physics based model with real world experimental results. This is the process that will be pursued in this section.

Following this process requires identifying and collating a set of experiment observations for the FFT performance of commercial CPU's. The major challenges in benchmarking this metric stems from fact that there are many different programs available to implement the FFT algorithm, and the performance of a given CPU will vary depending on which algorithm is chosen. An explanation of why this variation occurs would require a more elaborate digression into processing architectures that would beyond the scope of this work, but it is sufficient to state that the only consistent method of determining performance would be to evaluate hardware with a vendor-optimized program. For this reason, the most commonly cited benchmarking standard for FFT performance is the FFTW process develop Frigo and Johnson at the Massachusetts Institute of Technology [65]. Their approach has been shown to be competitive with vendor optimized programs, but is not tuned to a specific machine. Instead, FFTW uses a planning algorithm to adapt its algorithms to the hardware

www.manaraa.com

under evaluation in order to maximize performance. The input to this planner is a multidimensional loop of multidimensional DFT's, at which point the algorithm applies a rule set to decompose the problem into simpler sub-problems of the same type. Sub-problems at the lowest level of decomposition are then solved directly, using optimized code that is automatically generated by a special purpose compiler. In this way, different hardware can be evaluated against the same benchmarking standard.

Frigo and Johnson used their benchmarking algorithm to evaluate several hardware configurations from 200 to 2005, the results of which are available on line through the FTTW website[8]. Unfortunately, this dataset does not overlap with the SPEC dataset used in the previous sections, since the SPEC benchmarking standards were not finalized until 2006. The search for a replacement dataset proved to be difficult in the initial phases of this work, as the majority FFT experiments seemed to encompass, at best, a handful of platforms. Moreover, a significant number of these experiments did not provide the precise information necessary to identify the specific CPU being tested, which prevents the collation of results into a single, comprehensive database.

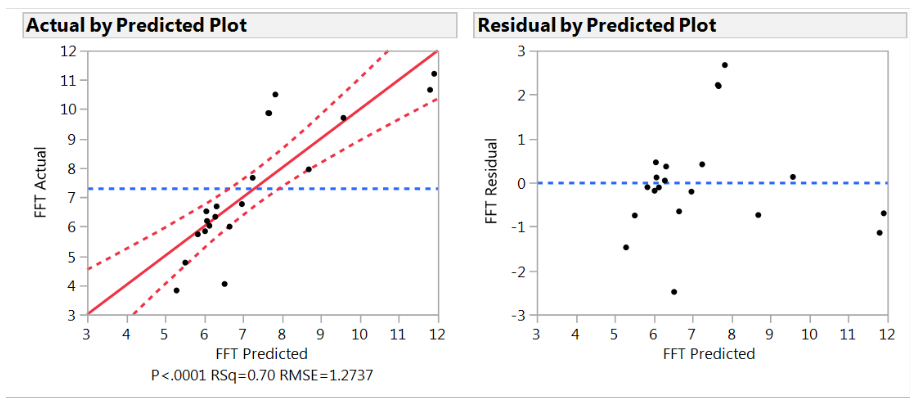
This scenario changed dramatically in January of 2016, when the FFTW algorithm was added to the Phoronix Test Suite. This benchmarking program allows researchers to upload their results to the open database at www.openbenchmarking.org, and the program automatically collects the necessary information to integrate results with other databases. Within the first two weeks of introduction, more than 100 independent FFTW entries were generated, and this volume of data was sufficient to formulate a surrogate performance model.

With the relevant data in hand, it is now possible to apply the model generation process described in the previous section. Figure 89 depicts the results of this process

[8]The interested reader is referred to the author's home site for further discussion: http://www.fftw.org/

المنارة للاستشارات
www.manaraa.com

for the standard RSM template. As the results show, there are two immediate problems with the baseline model. First, an R^2 value of 0.70 is well below the threshold of 0.90 necessary to validate the model. Second, only one of the metrics survived the iterative process of removing terms that were deemed to be statistically insignificant. This result is troubling, as it implies that the FFT performance of a CPU is only loosely correlated to the metrics defining the state of the art in this analysis.



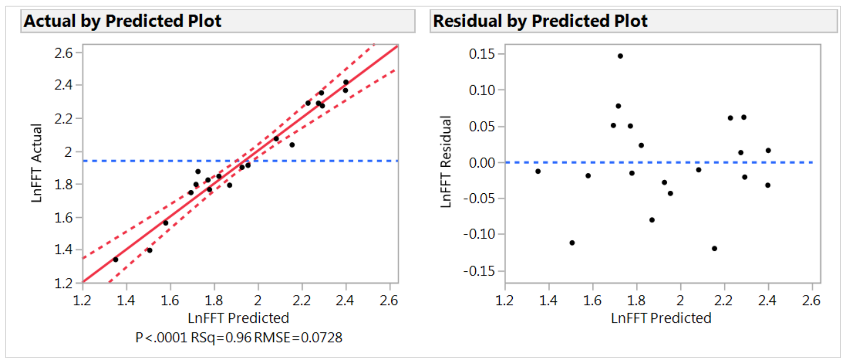
Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	4.4498	0.5258	8.46	<.0001
FPRate	0.0212	0.0033	6.46	<.0001

Figure 89: Initial Model Fit for Performance Surrogate

Having determined that the baseline RSM template is insufficient to model the problem, the next step in the model generation process stipulates that the analyst should consider either a transformation of the response variables or inclusion of higher order terms. As there is little evidence at this point to suggest that one approach would be preferable to another, both models were generated. Specifically, the transformed model replaced the original responses with their natural logarithm prior to fitting the model. The exponential function was then be applied to the output of this model in order to map responses from the logarithmic domain to the original performance domain. To generate the higher order model, the original RSM template was replaced with a full factorial template, which added third and fourth order terms. In addition, parameters failing to satisfy measures of statistical significance in either

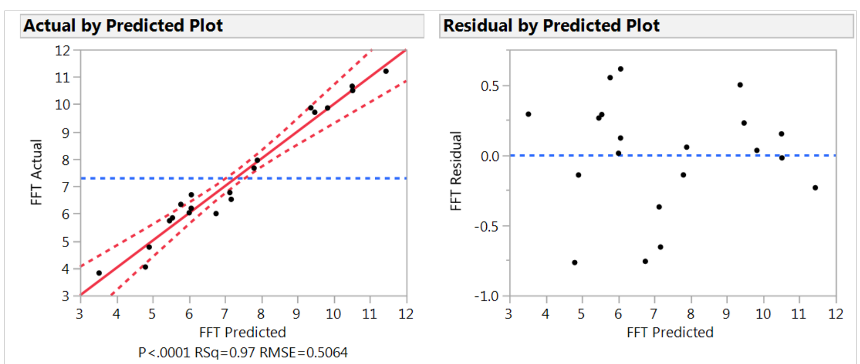
www.manaraa.com

model were removed using the same process applied to the original RSM template. Results of these analyses are provided in Figure 90 for the transformed model, and Figure 91 for the higher order model.



Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	1.2063	0.0444	27.18	<.0001
FLOPS	0.0097	0.0024	3.96	0.0013
FP	0.0108	0.0010	10.96	<.0001
FLOPS*FLOPS	-0.0005	0.0002	-3.04	0.0083
INTRate*INTRate	-2.76e-6	9.272e-7	-2.98	0.0094

Figure 90: Response Transform Fit for Performance Surrogate



Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	2.1569	0.5044	4.28	0.0011
FP*INTRate	0.0010	0.0002	4.53	0.0007
FLOPS*FP*INTRate	0.0004	6.323e-5	6.88	<.0001
FPRate	0.0396	0.0038	10.39	<.0001
FLOPS*FPRate	0.0011	0.0003	3.34	0.0059
FLOPS*FP*FPRate	-0.0003	6.772e-5	-5.38	0.0002
INTRate*FPRate	-0.0003	5.317e-5	-5.18	0.0002
FLOPS*INTRate*FPRate	-4.945e-5	7.184e-6	-6.88	<.0001

Figure 91: Full Factorial Fit for Performance Surrogate

At first glance, these fits both appear to be a substantial improvement over the original RSM template. Both models comfortably exceed the R^2 threshold of 0.95, their residuals contain no discernible trends, and all metrics used to fit the MDGM

المنارة للاستشارات
www.manaraa.com

are included. However, both models included highly non-linear terms to achieve their respective fits, and it is there for prudent to conduct another extrapolation study. The results of these studies are provided in figure 92 for the transformed fit, and Figure 93 for the full factorial fit.

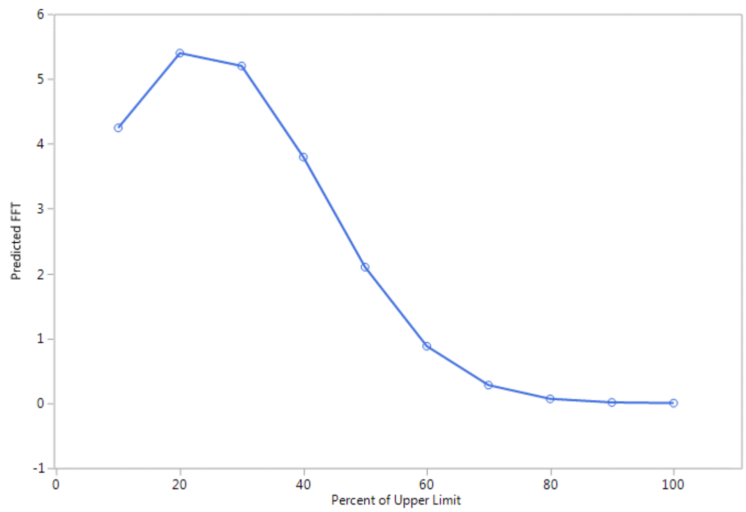


Figure 92: Extrapolation Study for Transformed Performance Surrogate

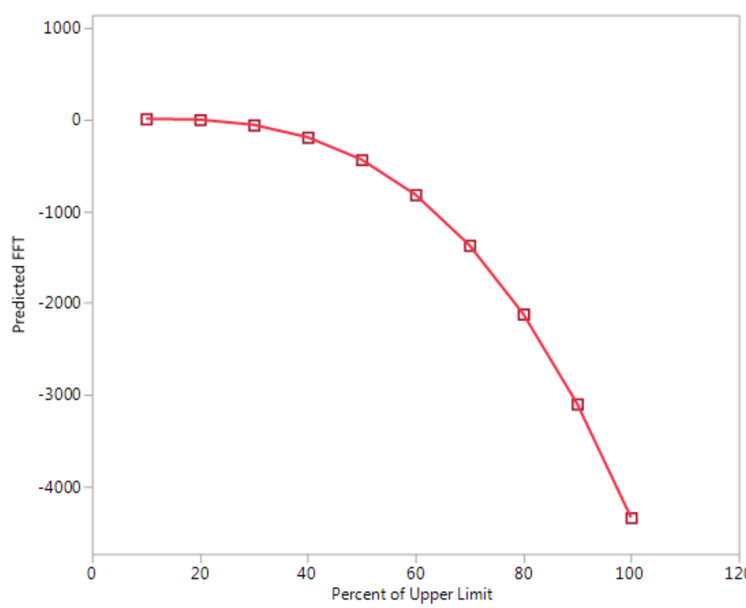


Figure 93: Extrapolation Study for Full Factorial Performance Surrogate

Clearly, the extrapolation tests demonstrate that both models are entirely inappropriate, even though they improve on the regression tests for suitability. The question

المنارة للاستشارات
www.manaraa.com

is why this would be true? The FFT algorithm is based on a series of floating point operations, and the SPEC FP and SPEC FP Rate metrics are explicitly formulated to capture the performance of a CPU in performing these operations. As such, one would expect that FFT performance would, at a minimum, be highly correlated with both metrics. This observation led to consideration of the underlying correlation structure between design and response variables as a possible factor influencing the fit process. Figure 94 provides the results of this correlation analysis.

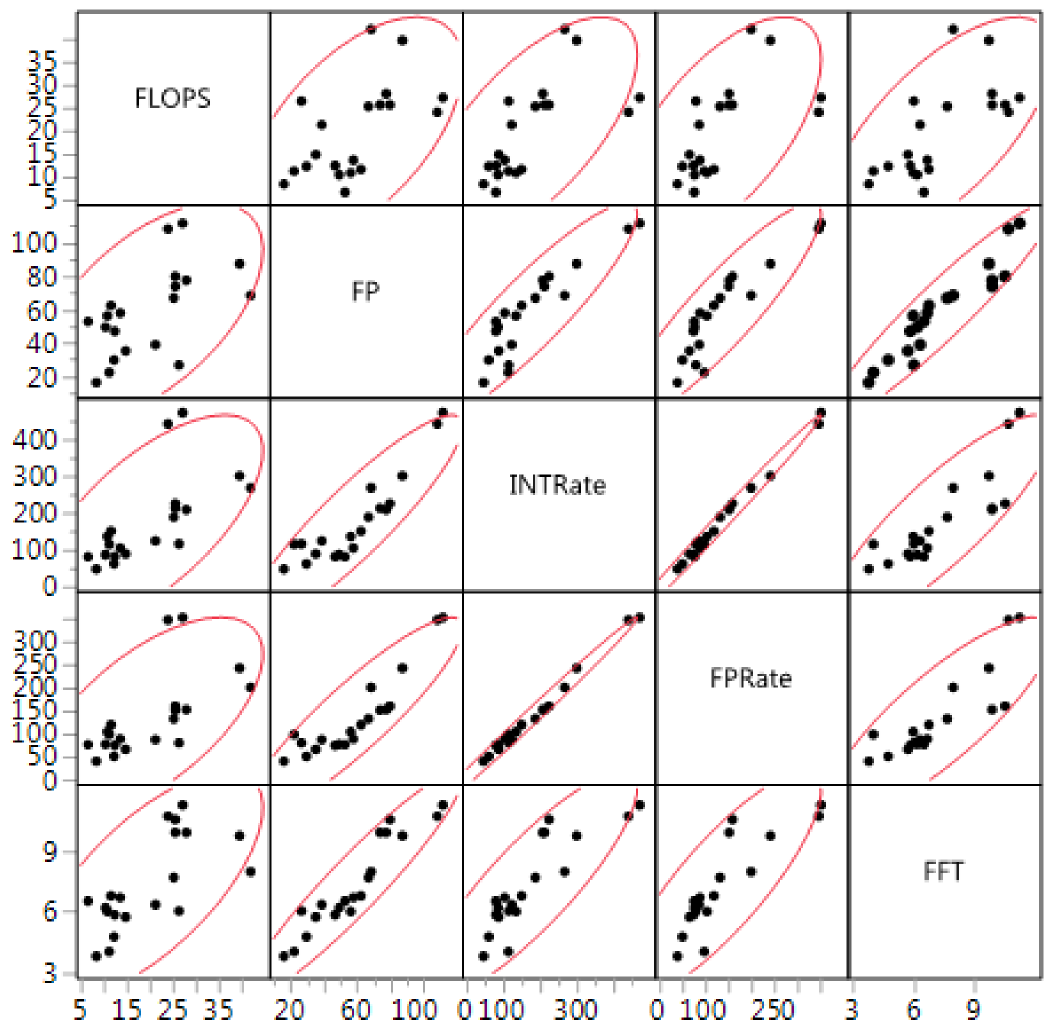
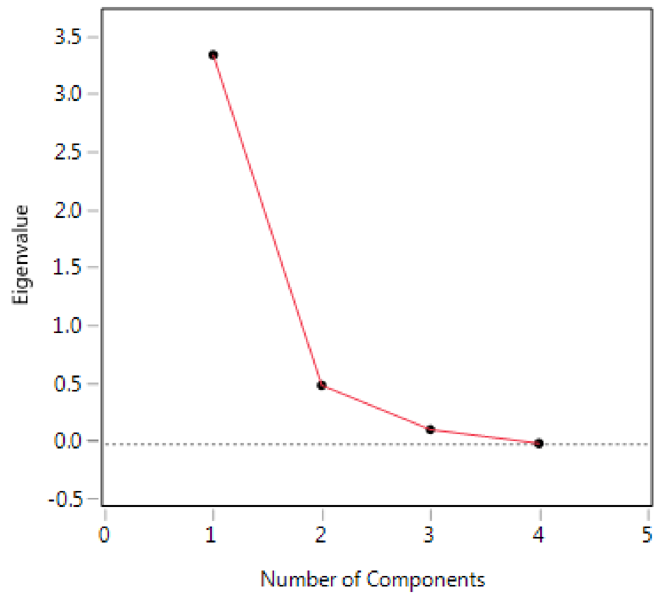


Figure 94: FFT Correlation Among Design and Response Variables

Figure 94 demonstrates, quite clearly, that there is in fact an incredibly strong, positive correlation between design and response variables. This demonstrates that

المنارة للاستشارات
www.manaraa.com

the first portion of the hypothesis is true, but it remains to be seen if this correlation is in fact the underlying feature complicating the fit process. The most direct way to evaluate the this proposition is to refit the model to the principle components of the design variables, as opposed fitting directly to the design variables. Figure 95 present the results of a Princple Component Analysis (PCA) on the design variables.



Eigenvalues

Number	Eigenvalue	Percent	20 40 60 80	Cum Percent
1	3.3660	84.149		84.149
2	0.5062	12.655		96.804
3	0.1238	3.094		99.898
4	0.0041	0.102		100.000

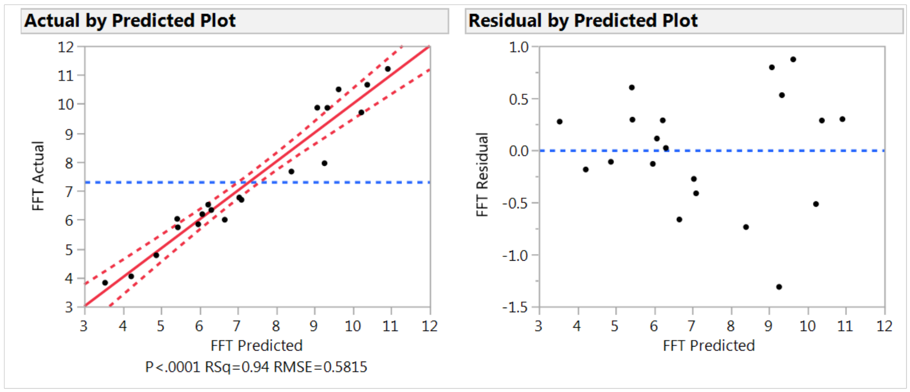
Eigenvectors

	Prin1	Prin2	Prin3	Prin4
FLOPS	0.41790	0.90071	0.10911	0.04650
FP	0.50721	-0.33198	0.79529	0.00601
INTRate	0.53527	-0.16208	-0.40356	-0.72412
FPRate	0.53064	-0.22854	-0.43902	0.68808

Figure 95: Design Variable PCA

The eigenvalues of Figure 95 indicate that 99.898% of the variability in design variables is explained by the first three principle components. The standard RSM template was therefore applied to those components, and the iterative process of eliminating statistically insignificant parameters was repeated. The fit results are provided in Figure 96.

المنارة للاستشارات
www.manaraa.com



Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	7.2389	0.1301	55.65	<.0001
Prin1	1.0318	0.0646	15.98	<.0001
Prin3	0.9194	0.2408	3.82	0.0014

Figure 96: Model Fit for Principle Component Surrogate

Here, the R^2 value is 0.93, which is well within the acceptable range, and the distribution of residuals shows no discernible patterns. It should also be noted that the final model is composed of only two linear terms, where each term is itself a linear combination of the four design variables. This corrects for the deficiency on the original RSM fit, where only one of the four design variables was utilized. As such, all variables necessary to define the state of the art at a given time are also necessary to define a component's FFT performance, which conforms to intuition. The final evaluation step is to then perform the now familiar extrapolation study on the model, the results of which are provided in Figure 97.

A cursory inspection indicates that the results have improved substantially, as the FFT model now demonstrates the monotonic property in which increasing component performance increases FFT speeds. In addition, the limiting case in which component design variables reach 100% of their respective limit states yields an FFT value of 18.78 GFLOPS, which is approximately 60% greater than the best score found within the historical database. This appears to be a reasonable result, which leads to the conclusion that the proposed model is sufficient to proceed.

المنارة للاستشارات
www.manaraa.com

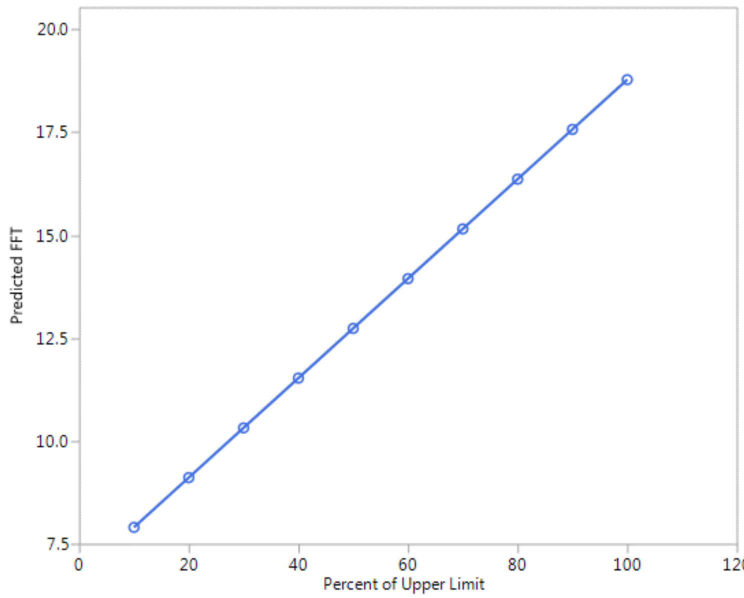


Figure 97: Extrapolation Study for Principle Component Surrogate

8.5 Step Five - Trade Space Exploration

8.5.1 Additional Modeling Assumptions

The worked performed in the previous sections provides all of the major tools necessary to proceed with a trade space exploration. Step One defined requirements for the MCA using the CRM development process described in Chapter Four. The performance profile associated with the CRM provides the first source of constraints for any potential TRP's, since the performance profile associated with the TRP must be greater than the requirement profile across the system lifecycle. Step Two related the CRM to a specific, modular component within the greater system architecture. Step three identified the performance metrics associated with that component, and fit a growth model to an historical database of COTS components produced over the last decade. This MDGM provides the second source of constraints in the trade space analysis, as any performance improvements believed to be possible at a future date must conform to the state of the art at that time. Finally, Step Four generated two surrogate models relating component design variables to both cost and performance.

With that in mind, the trade space exploration process defined by Figure 57 can

المنارة للاستشارات
www.manaraa.com

proceed as initially described. To review, the timing of technology infusions is controlled by the outer loop of the algorithm, which idealizes timing assumptions as a binary string: zero indicates that a technology infusion is "turned off" for that time period, and a one indicates that the infusion is "turned on". Those assumptions are then passed to the interior loop. In this phase, a given member of the current generation is defined by a vector of n-1 performance attributes, subject to the constraint that values associated with each design variable lie between the respective starting and limits points used to develop the MDGM. Equation 35 is then applied to determine what value the remaining design variable must take in order to be consistent with the state of the art at that time. The complete design vector is then passed to the cost and performance surrogates to estimate the product's cost and performance. This process is repeated across all time periods in which a technology infusion occurs, and the results are converted into an aggregated cost estimate and performance profile. Next, the performance profile is compared against the requirement profile to determine both the dynamic value and number of constraint violations generated by the TRP. These values are then returned to the NSGA-II algorithm to determine which members of the population proceed to the next generation. Repeated iterations of the interior loop converge to the true Pareto Frontier associated with the given timing assumptions. Finally, retaining only those TRP's shown to be efficient across all timing assumptions yields in the final trade space associated with the OSA.

This review indicates that some additional assumptions are required. First, Equation 35 does not operate on the timing of the technology infusion, but on the date at which technology development process begins. The two parameters are related by the time required to execute the development process, which is assumed to be 6 months in this scenario. In addition, it is further assumed that technology infusions will only be considered on an annual basis, a minimum of one year must elapse between successive infusions, and no upgrades will be permitted in the final three years of the life cycle.

المنارة للاستشارات
www.manaraa.com

These assumptions dictate that the timing string be composed of 17 characters, where the first and last characters represents 2016 and 2033 respectively. Finally, because the baseline system does not exist in this scenario, it must be designed in parallel with the TRP. This requires that the first character in the binary string remain fixed at 1 for all iterations of the outer loop.

The second consideration that must be addressed is the fact that the CER is calibrated to calculate the cost of a COTS CPU, whereas the system component being upgraded is the full SBC. Some relationship is therefore required to relate the two. In the vernacular of the acquisitions cost-estimation techniques, this would best be accomplished through an engineering build-up model. However, the information necessary to develop this type of model is not available for the scenario currently under consideration. Moreover, this type of model would simply represent a shift of component costs across the cost axis, but would likely not provide a significant impact to the broader trends that will be identified by the model. Without loss of generality, the simplified engineering build up model applied in this scenario assumes that 25% of the SBC is a fixed ruggedization cost, and that the remaining cost is proportional to the CPU cost. These parameters were then calibrated to a real world price quote from a major manufacturer of SBC's[9], and the result is given in Equation 76.

$$Cost_{SBC} = \$5,000 + 12.5 \cdot Cost_{CPU} \tag{76}$$

The last consideration that must be addressed is specific to this scenario. Whenever a technology infusion is applied, there is likely a fixed, Non-Recurring Engineering (NRE) cost associated with technology development and system recertification. In addition, a sustainment strategy is required to determine how many components will purchased, when those purchases occur, and how that those decisions will be impacted by the decision to infuse new technology. These two interrelated factors, NRE and

[9]The author was not authorized by the manufacturer to release the exact figures

المنارة للاستشارات
www.manaraa.com

purchase rate, can only be determined through a proper life cycle sustainment model. The proposed methodology assumed that this model exists in the earlier phases of development, but there is insufficient information in the current experimental scenario to create such a model[10]. A further simplification is therefore required. To that end, all results will be presented in terms of the time-weighted average of component costs. In other words, if one assumes zero NRE costs, which is a common assumption under the "plug-and-play" idealization, and a constant purchase rate, then multiplying the average cost by the total number of components purchased would yield the actual life-cycle cost. The exact calculation of this metric is presented in Equation 77. Here, avg is the average component cost, n is the number of upgrades, c_i is the cost of the i^{th} upgrade, T is the total system life cycle duration, and t_i is the length of time in which the system utilizes the t_i component. Again, this assumption is necessary to present the results of this analysis in a concise form, but it is strongly urged that a true sustainment model be applied for practical applications.

$$avg = \left(\frac{1}{T}\right) \sum_{i=1}^{n} t_i \cdot c_i \tag{77}$$

8.5.2 Evaluation of Results

Having established assumptions for the proposed scenario, the remaining task is to generate and evaluate the results. To that end, the first step in the analysis was to develop the trade space for a closed architecture in order to provide a baseline that can be used to assess the significance of open architecture TRP's. This initial analysis therefore assumes that the system is designed in 2016, and is not modified thereafter. Figure 98 presents the results of this analysis across 100 generations, with

[10] The proposed methodology is intended to function in parallel with a complete life cycle sustainment model, but the complexity of developing a model to consider these factors is far from trivial. This area is the focus of a second thesis at the Georgia Institute of Technology, and will therefore not be addressed in a rigorous fashion

www.manaraa.com

50 members per generation and 20 binary characters per design variable.

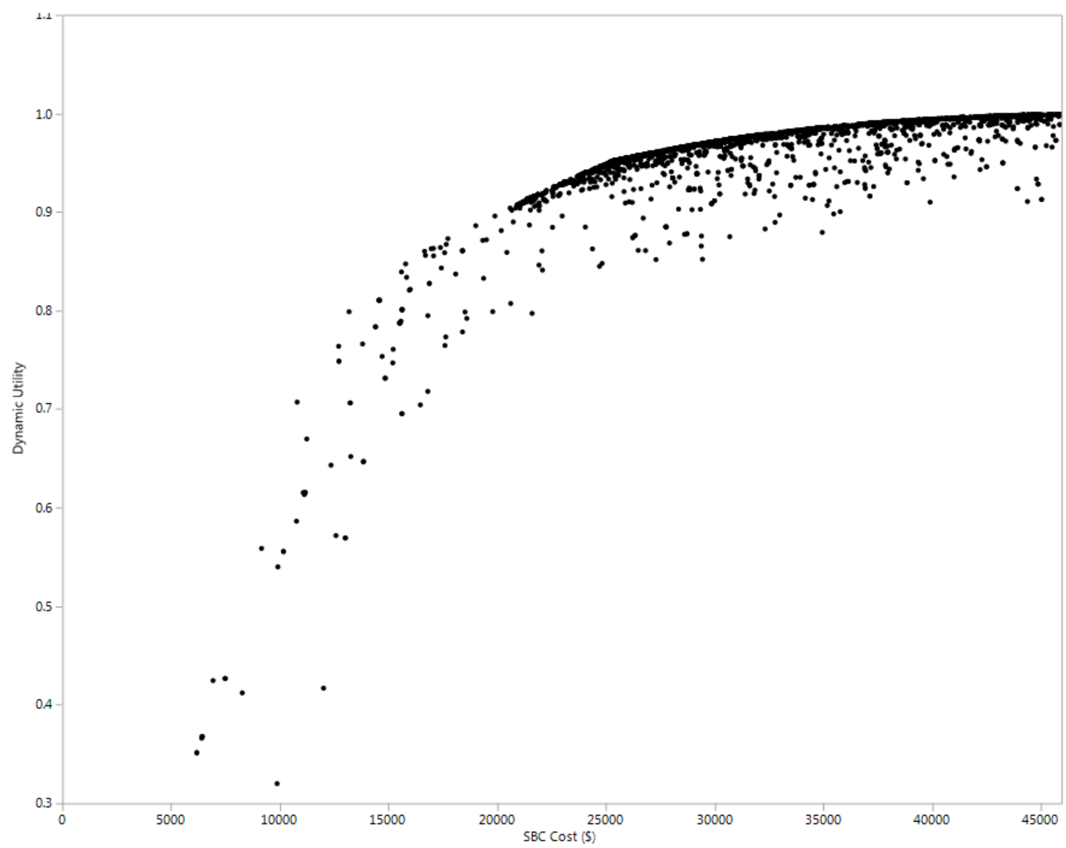


Figure 98: Raw Trade Space for No Upgrades

Considering the basic features of Figure 98, it appears that the trade space surface is fairly well defined, and possesses a reasonable orientation toward the utopia point of low cost and high performance. These qualities provide substantial support to the validation argument that the proposed trade space exploration process performs as intended. Moreover, it would appear that building the system today is quite an attractive alternative, as there numerous alternatives available at low cost that appear to provide surprisingly high performance. However, it is rather curious that there appears to be a much higher density of points on the upper portion of the frontier, where cost and performance are both high.

Figures 99 and 100 on the following pages provide an clear answer to this mystery.

Figures on the preceding pages should be interpreted in the following manner.

المنارة للاستشارات
www.manaraa.com

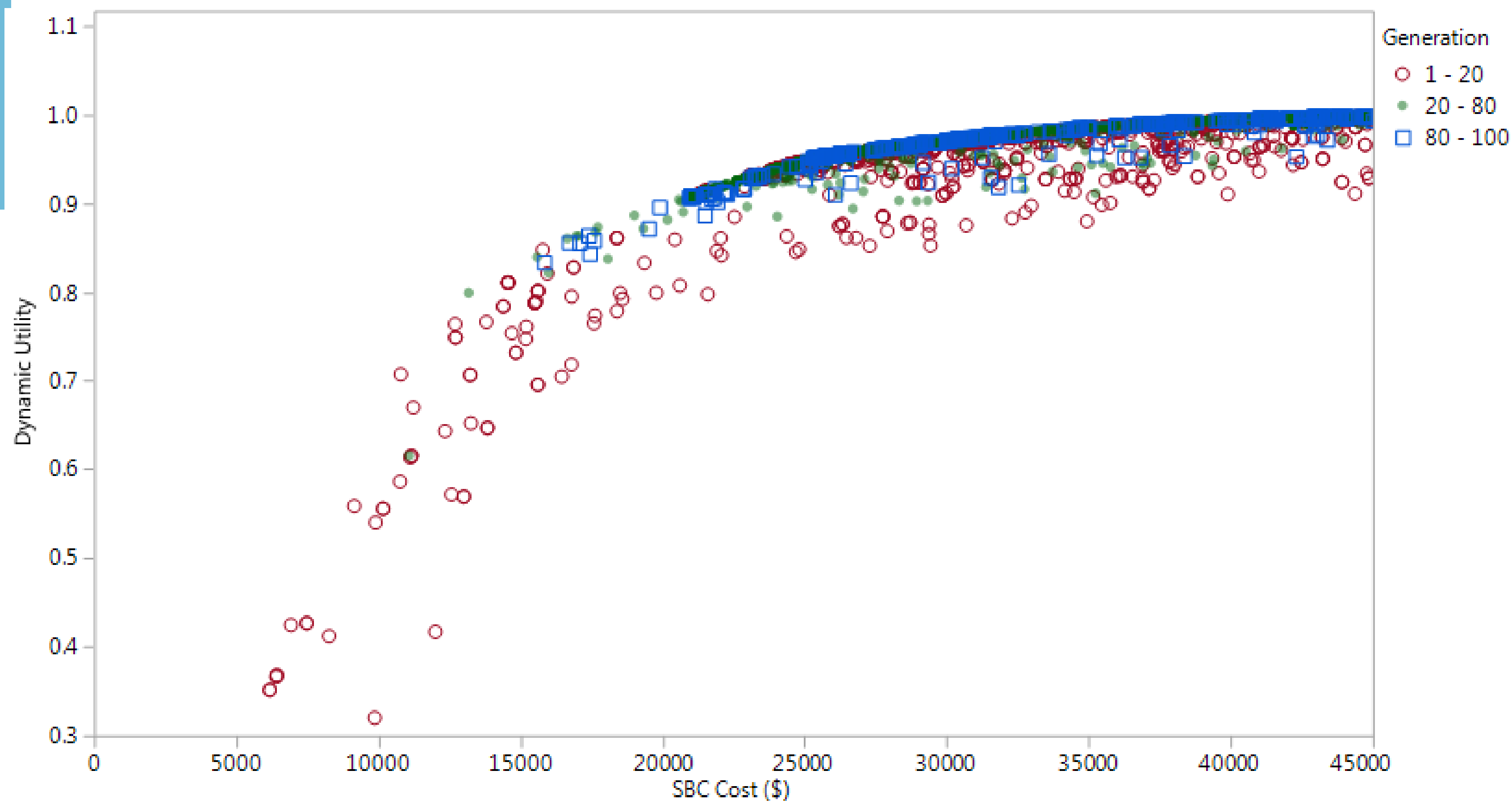


Figure 99: Progression of Generations During Trade Space Exploration

www.manaraa.com

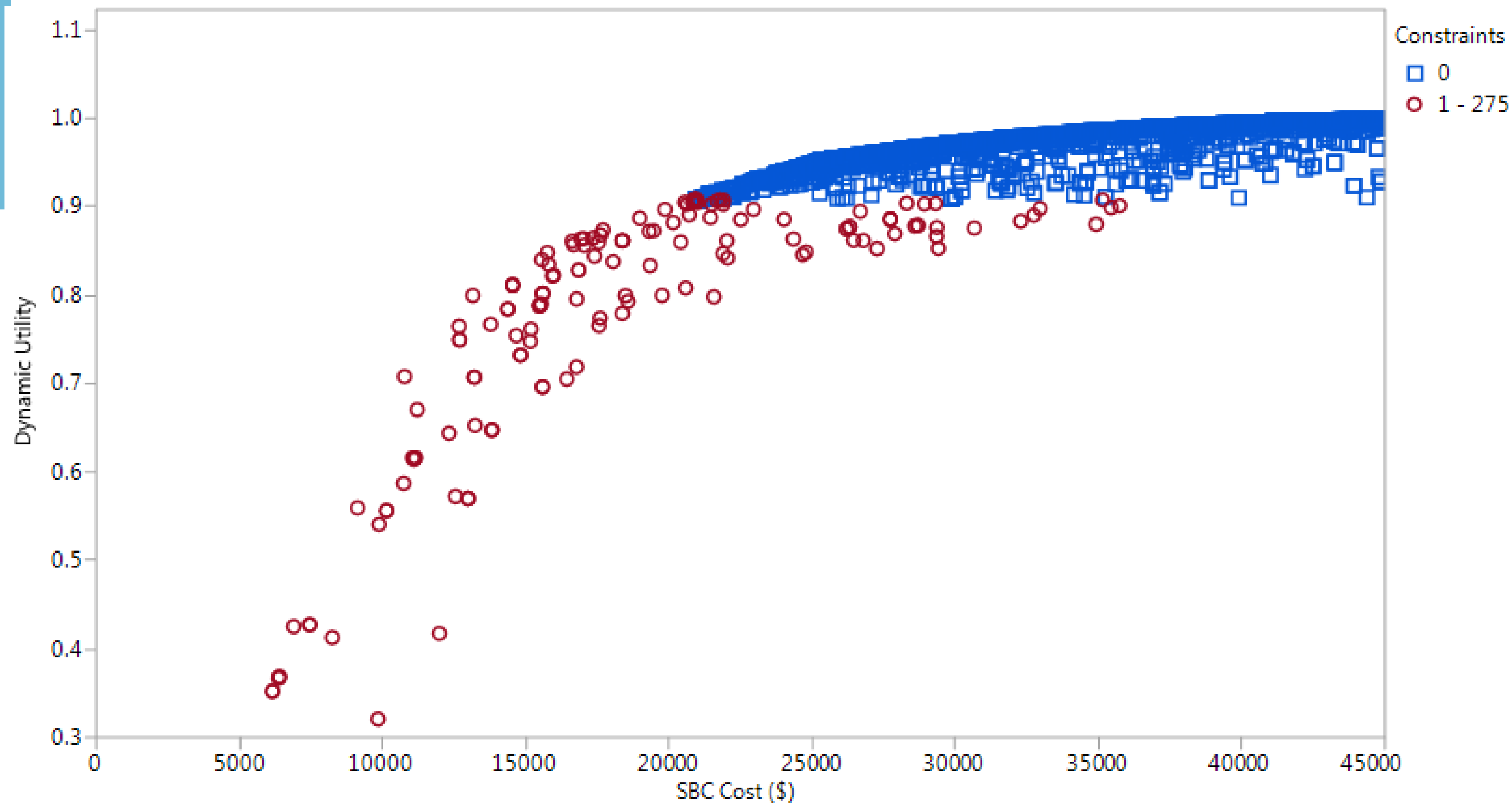


Figure 100: Constraint Violation Among Closed Architecture Alternatives

للاستشارات المنارة
www.manaraa.com

Figure 99 tracks the progression of the NSGA-II population across multiple generations, with the red circles indicating the first 20 generations and the blue squares representing the last 20 generations. Many of the points that appear to form the lower portion of trade space (i.e. low cost, low performance) lie in the early generations, but were immediately abandoned for points in the upper portion (i.e. high cost, high performance) of the frontier. Figure 100 makes it quite clear why this occurs - constraint violations. Specifically, the red circles indicate designs with one or more instances in which the performance profile failed to accommodate the scenario's requirement profile. Blue squares, on the other hand, represent those points that did satisfy constraints at all times[11]. The conclusion from these observations is that the optimizer drove early generations to the only viable portion of the frontier. This is precisely how the NSGA-II algorithm is intended to function, which provides further evidence that trade exploration process functions as intended.

Figures 99 and 100 demonstrate another important result. Recall that the initial theory of gold plating in acquisition programs is essentially a self-fulfilling prophecy. Systems are expensive and difficult to modify, therefore requirements must be derived from the end of the system's life cycle in order to ensure that it remains operationally effective. However, imposing excessive requirements drives the system to the extremes of the design space, which necessitates a complex, highly integrated design. Such designs are inherently expensive and difficult to modify, thereby perpetuating the cycle. The results expressed in Figures 99 and 100 almost perfectly articulate the second portion of this logic: the requirements imposed on the system quite literally drove the exploration process to the extremes of the design space. The author could find no other empirical observation of this commonly cited trend in OSA design, which implies that the significance of these results go beyond mere model validation.

[11]Constraint violations were evaluated at 500 discrete points along the system life cycle, and the number of violations in the legend of Figure 100 refers to the number of points where the constraint was not satisfied. This value was suppressed for clarity.

www.manaraa.com

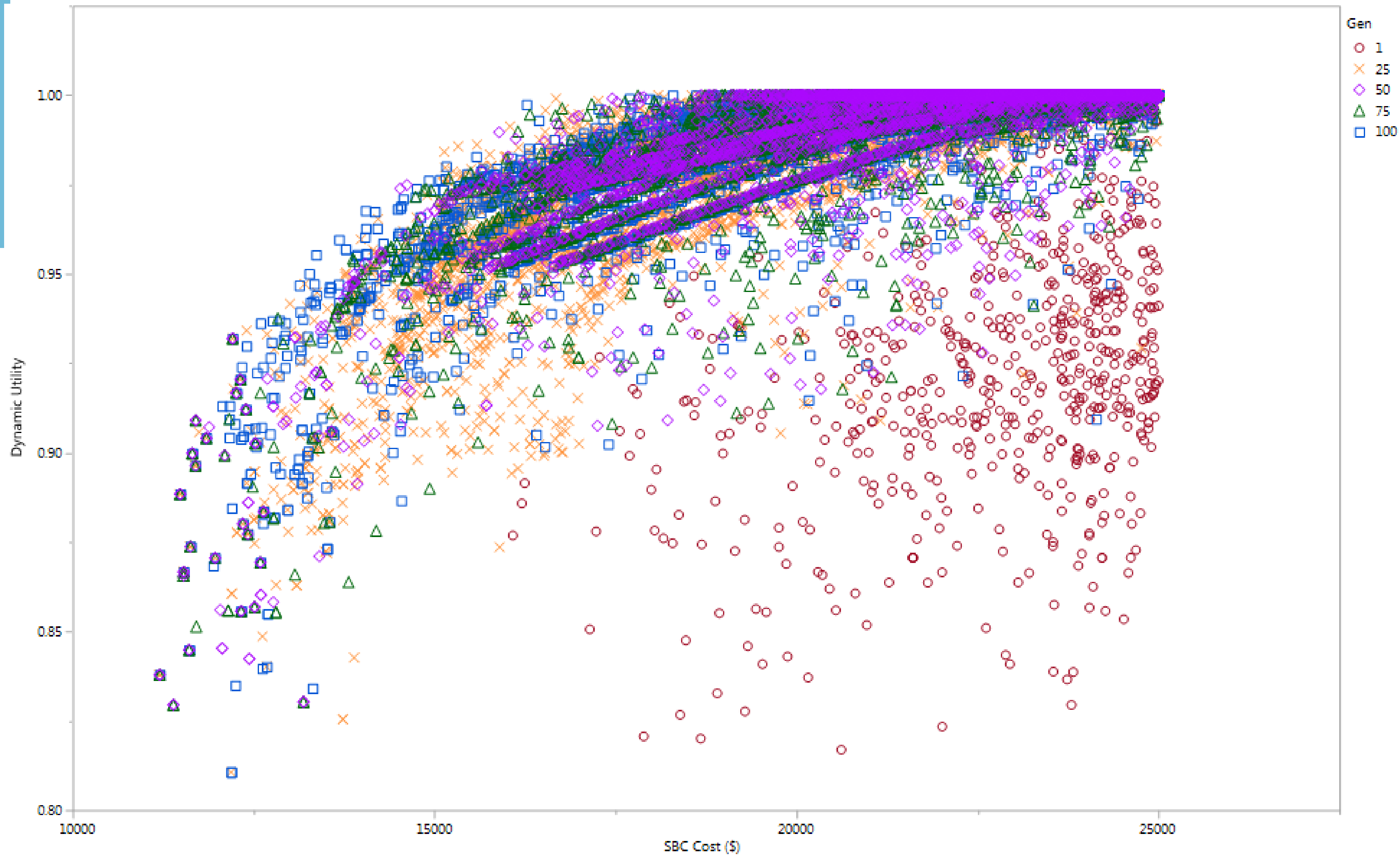


Figure 101: Trade Space Exploration for Two Upgrades

www.manaraa.com

Having established a baseline and validated the mechanics of the trade space exploration algorithm, the final step is to repeat the process across TRP's containing multiple upgrades. Detailed results of the exploration process for two upgrades are presented in Figure 101 in order to provide greater insight on the level of complexity in the trade space exploration associated with increasing the number of upgrades under consideration[12] For the sake of brevity and clarity, however, these results are not presented for all analyses performed.

It should be noted, however, that the exploration process is not intended to retain all of the data provided in the previous figures. Rather, the process is meant to progressively refine the frontier with each iteration of the outer loop (i.e. timing assumptions), such that the output is the efficient trade space for TRP's satisfying the scenario's expected performance profile. Applying this iterative filtering scheme leads to the final result provided in Figure 102, which presents the trade space associated with TRP's containing zero (i.e. closed architecture), one, two, or three upgrades.

One final consideration remains to be addressed. The multiple curves depicted in Figure 102 exist due to the lack of a comprehensive sustainment model, which was extensively covered at the start of the section. If such a model were present, then progressively increasing the number of upgrades would be penalized by the requirement to pay NRE costs multiple times. On the other hand, the lower unit costs associated with more frequent upgrades would provide greater cost savings as the number of units purchased increases. This balance would provide an interesting case study for future work, but the direct consequence is that all TRP's would be evaluated in terms of TLCC as opposed to average unit costs. This allows the multiple curves present in Figure 102 to be reduced to a single, efficient trade space for further consideration. This is the trade space that would inform the final step of the methodology.

[12]All points in Figure 101 were filtered to remove any TRP's that failed to meet constraints. This provides a more succinct picture of the exploration process.

المنارة للاستشارات
www.manaraa.com

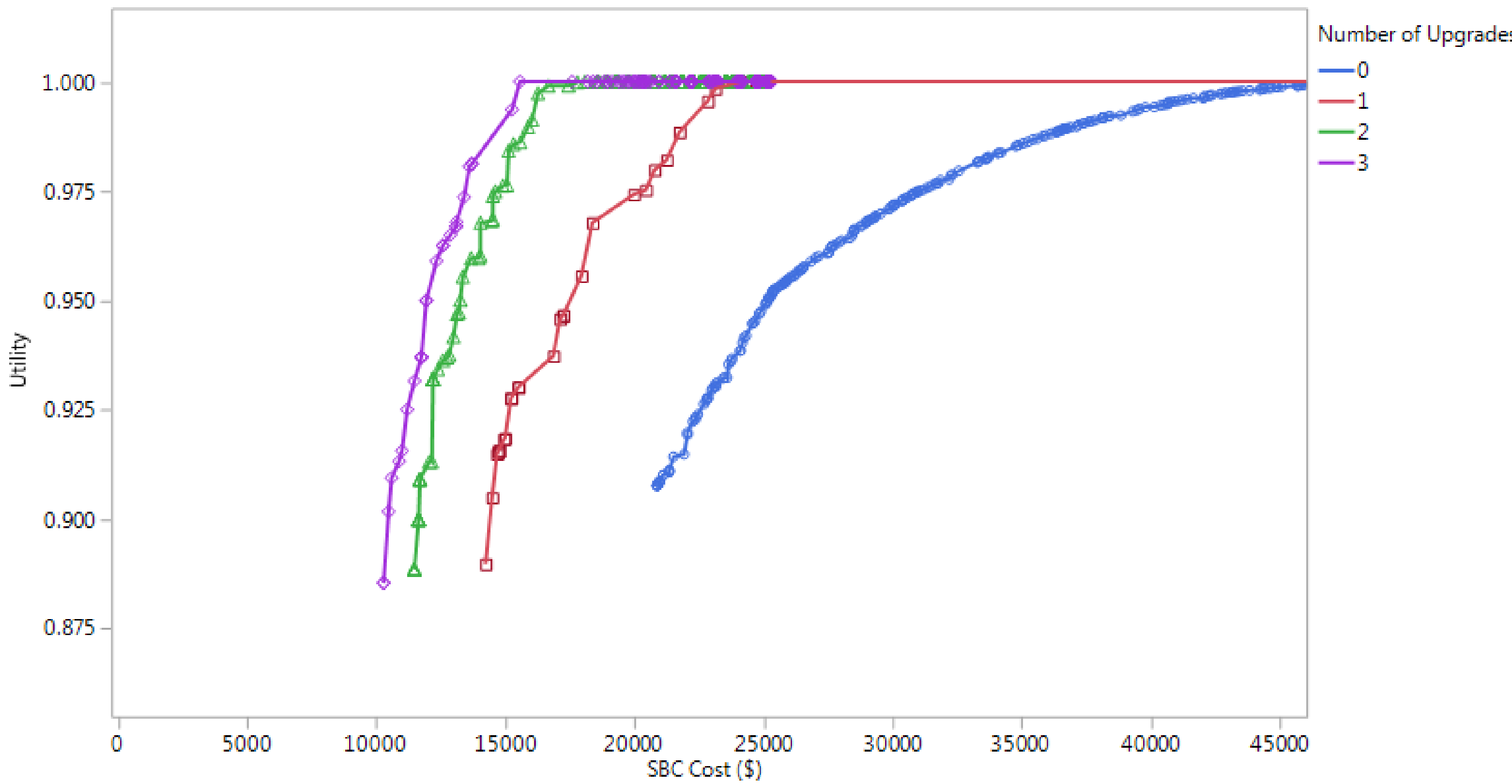


Figure 102: Combined Trade Space for Multiple Upgrades

www.manaraa.com

8.5.3 Identifying Families of Solutions

Chapter Five argued that the use of performance profiles to describe refresh plans is necessary to formulate a well posed problem, but the description is insufficient for decision making purposes. This statement followed from two observations. First, it is not clear what trades the decision maker should consider when comparing alternative profiles (e.g. is it better to have less capability now, or more capability later). Second, even if the decision-maker could articulate their preferences, the superposition of numerous performance profiles on the same axes would quickly obscure these trades. This latter observation was originally demonstrated with a simple cartoon, but results from the previous section now provide experimental support for the original statement. For example, the exploration process for a single upgrade identified 15,904 feasible refresh plans. If the performance profiles were used as the sole decision making criterion, then Figure 103 would serve as the trade space visualization for this scenario.

As predicted, the sheer volume of feasible plans prevents any serious evaluation of the merits associated with the cost and benefits of pursuing one plan over another. Such a formulation is clearly not an acceptable decision support framework, but it does provide a final point of validation for the modeling process. Specifically, each of the feasible profiles generated during the exploration phase is above the threshold requirement profile, depicted by the solid black line, at all points in time, which indicates that the system will always be able to perform its intended mission. In addition, there are only a handful of outliers with a starting performance at or below the level required to accommodate the first capability increment. This implies that both the present day requirements and those of the first capability increment can be reasonably met with existing technology. Therefore, the key decision point relates to how the refresh strategy will deal with the second and third increments, but no further information can be gained as to how this decision should be approached.

المنارة للاستشارات
www.manaraa.com

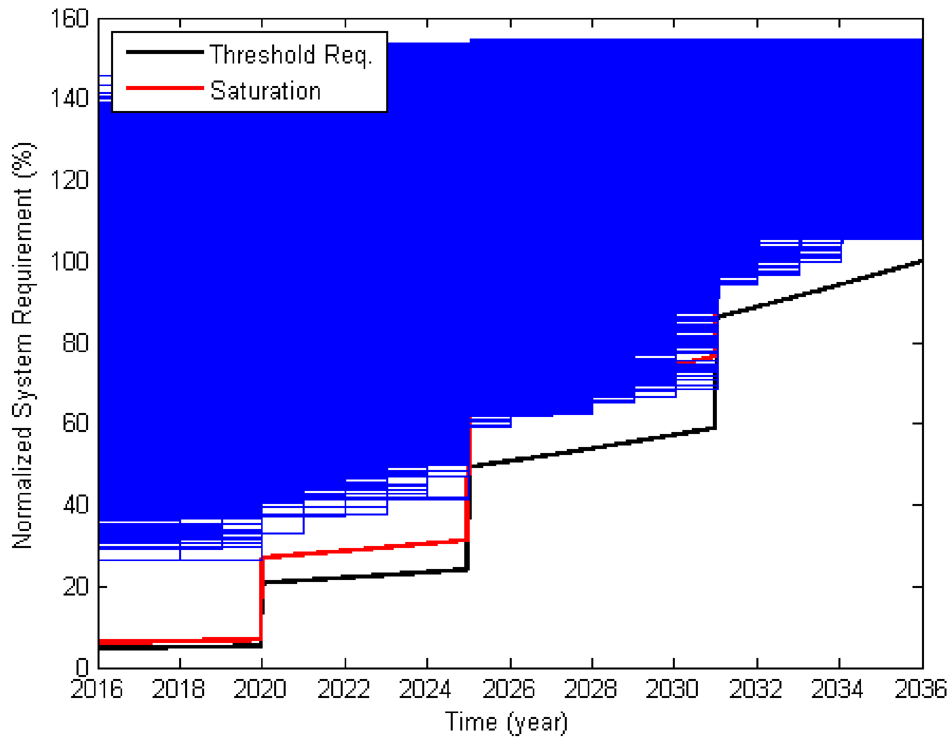


Figure 103: Decision Space Without the Application of Dynamic Value

Filtering the alternative plans with respect to dynamic performance and cost results in a much smaller set of efficient alternatives. A final point of consideration during this phase of analysis is whether the profiles of these efficient plans can be leveraged to gain further insight into the trades concealed by the scalar dynamic value score. To that end, consider again the efficient refresh plans for a single upgrade, and expand the definition of efficiency to include those plans that are close enough to the frontier to fall within the margin of error. This definition would result in the expanded trade space depicted in Figure 104, which, for the sake of clarity in the coming discussion, is partition into three sub-regions: low, moderate, and high cost/performance. Figure 105 presents the performance profiles associated with each of the points along this frontier.

المنارة للاستشارات
www.manaraa.com

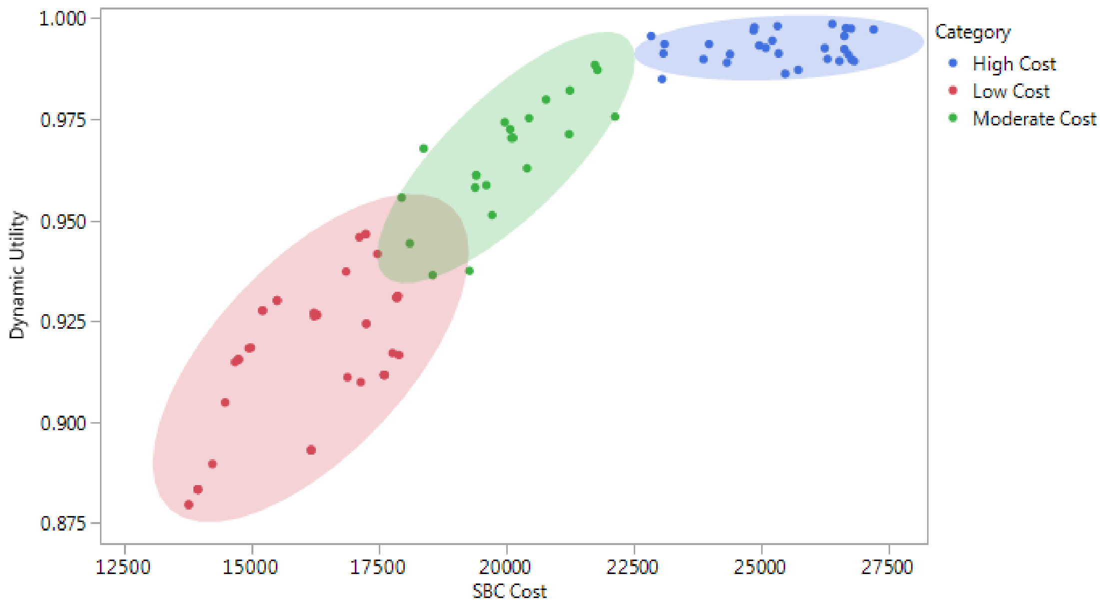


Figure 104: Decomposition of the Efficient Frontier Assuming One Upgrade

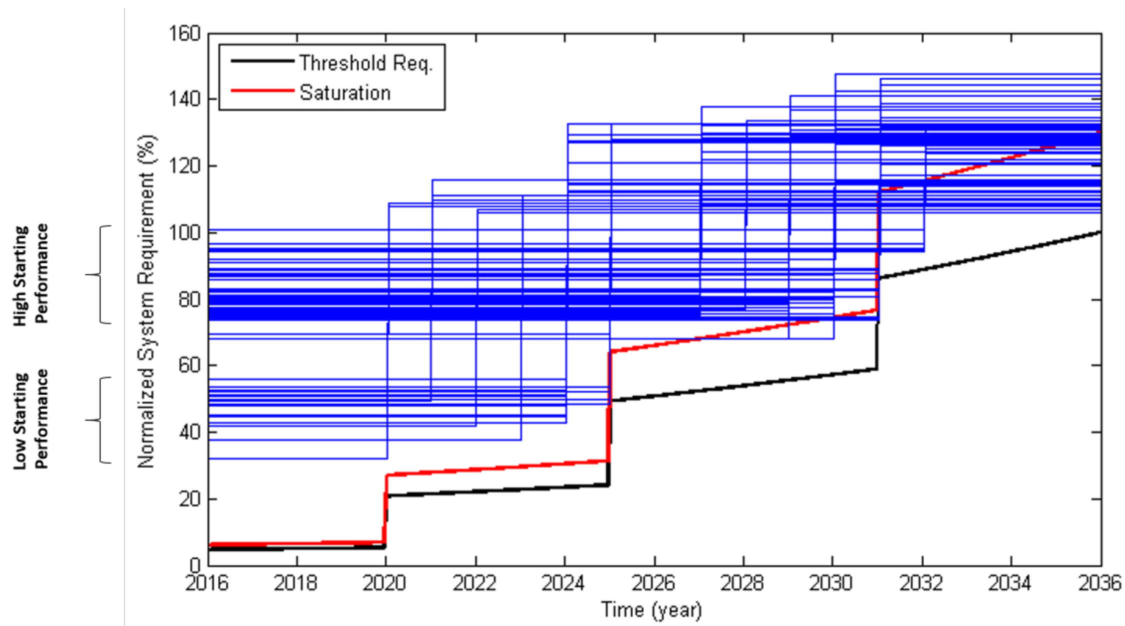


Figure 105: Performance Profiles for Efficient Refresh Plans

www.manaraa.com

The reduced set of profiles evaluated in Figure 105 is still somewhat difficult to interpret, but closer scrutiny does yield discernible trends. In particular, there appears to be a clear partition in terms of the starting performance of the system: roughly two thirds of the profiles start with relative high performance (between 75% and 95% of the terminal requirement), while the remainder start with relatively low performance (between 35% and 55% of the terminal requirement). Figures 106 and Figure 107 were constructed to isolate these distinct sub-groups of performance profiles in order to further resolve any underlying trends. In addition, the profiles in each figure were colored coded to indicate which region of Figure 104 the profiles map to under the cost and dynamic value transformations[13].

Figure 106 depicts the family of profiles corresponding to a low starting performance. First, note that all upgrades occur in a window between the first and second capability increments, which indicates that a modest initial performance will necessitate an early upgrade. This is not, in and of itself, a significant observation, since the threshold requirement is expected to grow at an increasing rate over time. The more substantive observations comes from the distribution of upgrade timings that lead to different regions of the partitioned frontier. Specifically, upgrades leading to the low cost region (indicated in red) are distributed between 2020 and 2024, whereas high performance upgrades (indicated in blue) occur between 2025 and 2026. It is also worth noting that the majority of alternatives in this family are contained within the low cost region of Figure 104, and there is only one alternative leading to the moderate cost/performance region (indicated in green). This implies that the low starting performance strategy provides a fair degree of flexibility in targeting the low cost region of the frontier, but this flexibility does not extended to the moderate and high cost regions. If these regions are desirable, then it is clear that the system and its

[13] A handful of outliers were omitted from Figures 106 and 107 in order to highlight the broader trends

المنارة للاستشارات
www.manaraa.com

upgrade should be formulated in a manner that is consistent with the complimentary family of alternatives.

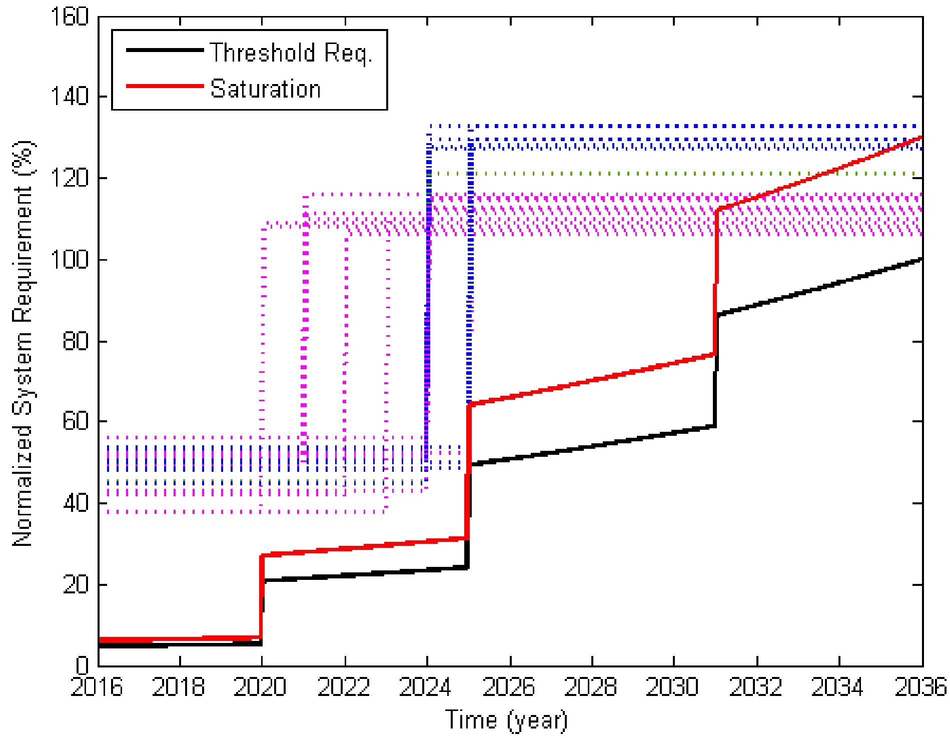


Figure 106: Efficient Profiles for Systems with Modest Initial Performance

Figure 107 depicts the complimentary, high starting performance family of profiles. The vast majority of upgrades in this group occur between 2027 and 2031, which is the same duration as the window of opportunity afforded by the low-cost alternatives in Figure 106 (i.e. 2020-2024 vs. 2027-2031). Unlike the previous family, however, the timing of upgrades does not neatly segregate profiles into disjoint sets targeting specific regions of Figure 104. Rather, it is possible to efficiently target any region of the frontier within this window by altering decisions regarding the cost/performance of the corresponding upgrade.

In conclusion, the underlying trend in efficient alternatives for this scenario equates to two distinct families of profiles. The first family is defined by a low starting performance that is significantly improved by a relatively early upgrade. This approach

المنارة للاستشارات
www.manaraa.com

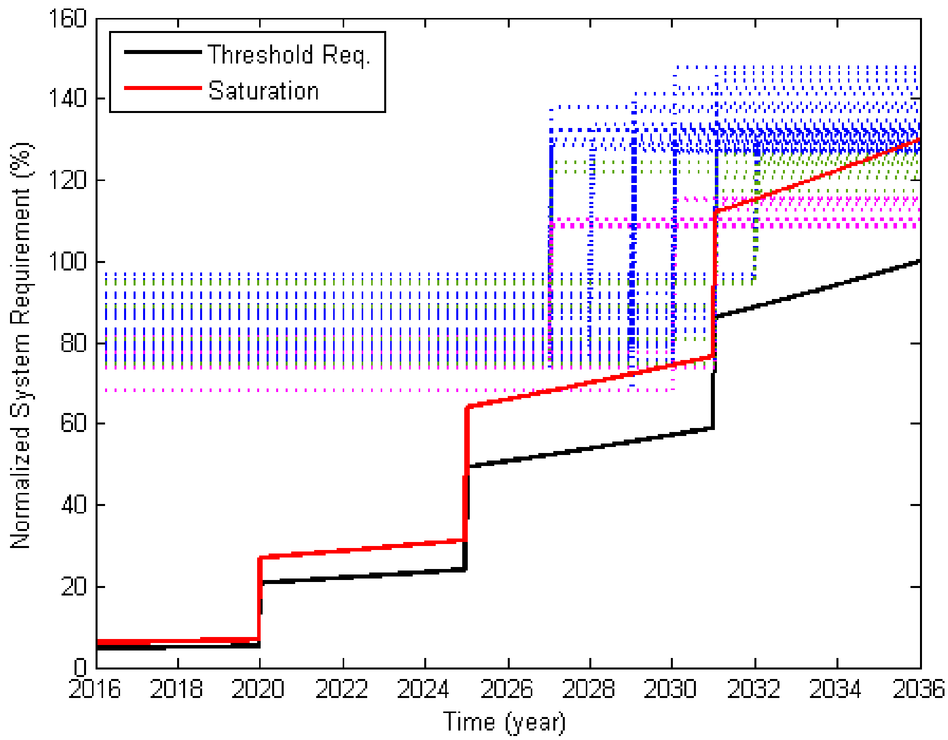


Figure 107: Efficient Profiles for Systems with High Initial Performance

reduces the initial development cost of the system, but provides few paths to target the higher cost/performance regions of the dynamic trade space. Conversely, the second family is defined by a system with a high starting performance that will be augmented with a much later upgrade. This strategy provides substantially greater flexibility to target alternative regions of the dynamic trade space, but it comes at the expense of increased development costs. These details were initially conceal by the simplified, dynamic value representation of alternative refresh plans, which implies that some measure of iteration between performance profiles and their dynamic representation can help inform the decision-making process.

المنارة للاستشارات
www.manaraa.com

8.6 Step Six - Sensitivity Analysis

8.6.1 Selection of Refresh Strategies

The results presented in Figure 102 clearly indicate that there is a substantial advantage in upgrading the MCA's SBC to keep pace with evolving requirements. Yet, the marginal value of this advantage declines sharply as more upgrades are added to the sub-system's TRP. The fundamental question to be addressed at this stage is whether that marginal reduction in component cost is sufficient to offset the additional NRE expenses. Unfortunately, addressing this question requires a rigorous sustainment model that is not available for the current analysis. It will therefore be assumed that a single upgrade provides the best value for the experimental scenario, and that the corresponding frontier serves as the final trade space presented to decision-makers.

The intent of this phase in the analysis is to have decision-makers review the trade space, and select a series of design points that should be subjected to further scrutiny. This scenario assumes that two points are selected: a minimum cost alternative, and a cost effective alternative. Figure 108 defines these points on the normalized scale that will be used for subsequent analyses.

Once identified, the next step is to use these points as a means of quantifying the decison-maker's preferences *a posteriori*. To accomplish this, it is first necessary to select an objective function with weight parameters that can be used to model these preferences. The specific model used in this analysis is the classical Overall Evaluation Criterion (OEC) given in Equation 78, where x and y represent cost and performance.

$$V = w_{cost}\left(\frac{x_i - x_{max}}{x_{min} - x_{max}}\right) + w_{perf}\left(\frac{y_i - y_{min}}{y_{max} - y_{min}}\right) \tag{78}$$

Note that the normalization parameters are inverted for cost and performance, since the two metrics have opposing quality characteristics: larger is better for performance, and smaller is better for cost. This is a necessary transformation to determine

المنارة للاستشارات
www.manaraa.com

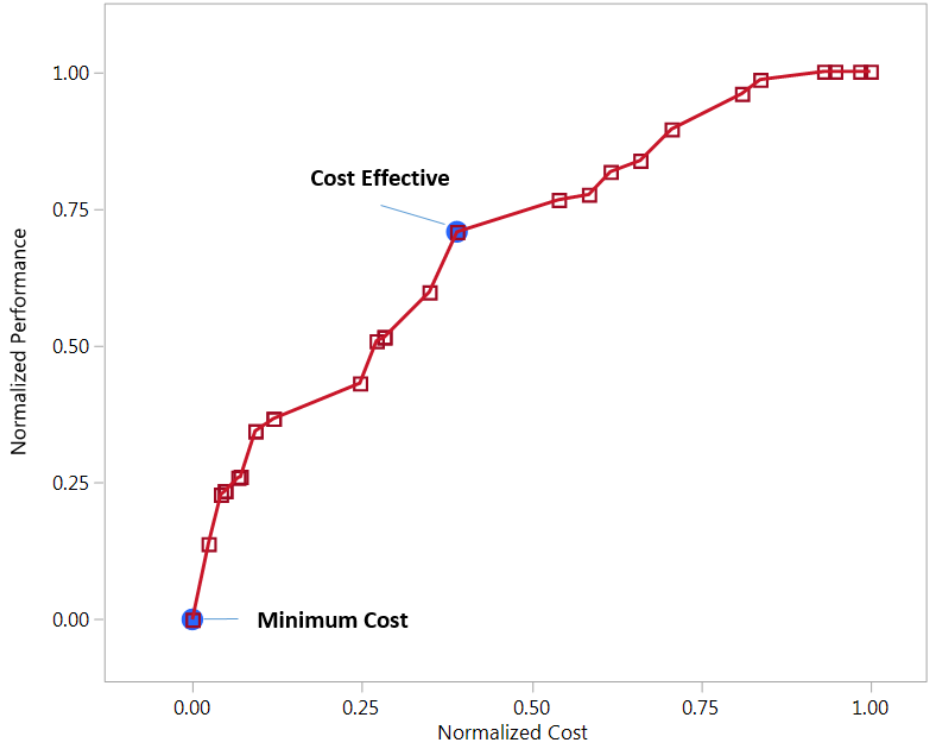


Figure 108: Identification of Points For Sensitivity Analysis

the proper weight parameters, but presenting results in this fashion would reflect the trade space across the cost axis. To maintain a consistent visualization, all results will continue to be depicted under the normalization scheme expressed in Figure 108, though numerical analysis will follow the functional form provided by Equation 78.

The final step in this process is to then determine which weight parameters in the OEC lead to the design points identified by decision-makers. This analysis is somewhat trivial in the case of minimum cost, where the weights are, by definition, 0 for performance and 1 for cost. The cost-effective point presents a less obvious case, and therefore requires the optimization method described in Chapter Six. The results of this approach indicate that the appropriate weights are 0.5238 for cost and 0.4762 for performance. Constant value contours for this function are depicted on Figure 109, and it is clear that the tangent point coincides with the prescribed point. This indicates that the objectives functions are validated for the next phase of the

www.manaraa.com

analysis.

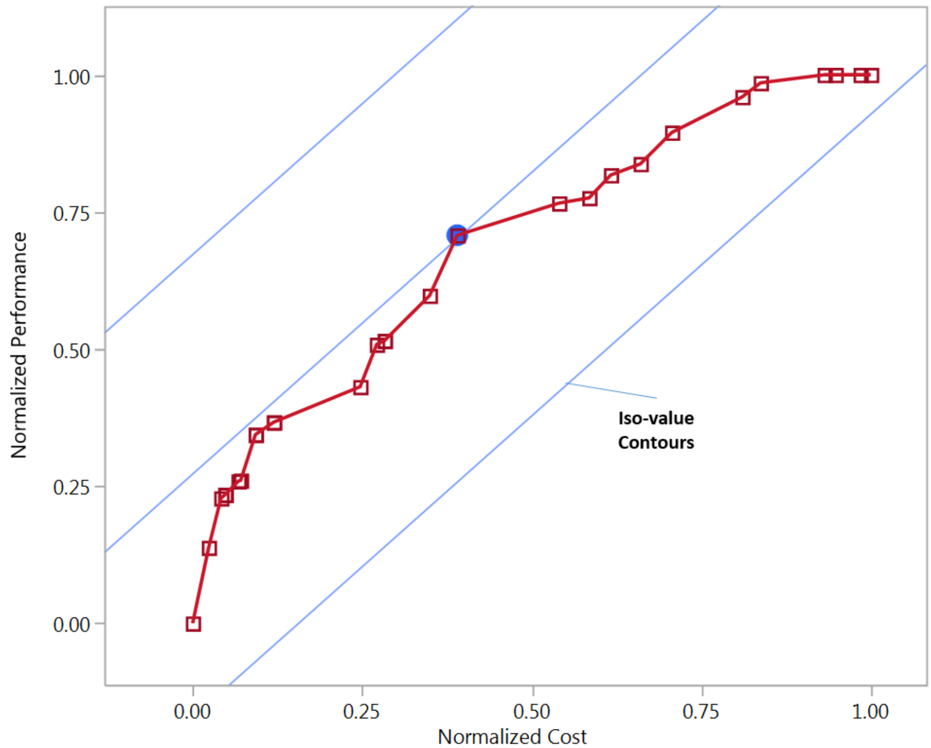


Figure 109: Validation of OEC For Sensitivity Analysis

8.6.2 Initial Assessment - Formulation and Validation

With an OEC available, it is now possible to conduct a MCS on the model to assess the sensitivity of design points identified by decision-makers to various drivers of uncertainty highlighted throughout the scenario development. Recall that the drivers of uncertainty inherent to the CRM were previously assumed to be normally distributed according to the parameters identified in Table 12. In addition, it was observed that the engineering limits associated with the various design variables and the rate of cost reduction in commercial technology are also subject to uncertainty. As there was no apparent logic by which to select an appropriate distribution, the principle of maximum entropy would dictate that the uniform distribution is the most appropriate

www.manaraa.com

model[14]. Consequently, it is assumed that all remaining parameters are uniformly distributed according to values given in Table 14.

Table 14: Left and Right Limits of Uniformly Distributed Uncertainty

	Left Limit	Right Limit
Theoretical FLOPS Limit	97.5	112.5
SPEC FP Limit	156	180
SPEC INT Rate Limit	1625	1875
SPEC FP Rate Limit	1105	1275
Annual Cost Reduction	1%	3%

From this point, the proposed methodology advocates that a bi-level, continuous optimizer can replace the more complicated trade space exploration. To review, the top level of the optimizer varies the timing of technology infusions, while the interior optimizer varies design variables in order to optimize the OEC developed in the previous section. To validate this approach, an initial study was performed in which the proposed optimization scheme was applied to the problem of finding the minimum cost for a configuration with no upgrades. This approach is advantageous in that it allows the interior optimizer to be directly evaluated, since there are no timing considerations for the closed architecture. In addition, no measures of uncertainty were introduced in this analysis so that the correct answer would be known. If the proposed approach functions as intended, then the optimizer should arrive at this correct answer from a variety of starting points.

Based on the logic established up to this point, a two step validation test was conducted. First, the optimizer was initialized with a vector of design variables that was identical to the design that provided the lowest cost alternative for a closed architecture. The value returned by the optimizer was $20,644.91, which is only

[14]The expected limits of design variables were developed under the assumption that the best values in the historical database could be increased by 40%. Left and right limits of the associated uniform distributions of uncertainty coincide, respectively, with an assumption of 30% and 50%

المنارة للاستشارات
www.manaraa.com

1.15% less than the value of $20,886.56 found in the trade space exploration process. This provides some initial support for the proposed MCS approach, but it provides significant validation for the exploration process. The second step in this process was to vary the initial design vector by a scalar factor in order to determine if the optimizer could arrive at the same conclusion for a farther start point. This test proved to be an abject failure. In short, if the initial start point is varied by more than 5%, then the optimizer is unable to satisfy the constraints and fails to converge. This level of performance is unacceptable, and a new approach is therefore required.

8.6.3 Stochastic Approach

There are several possible explanations for the failure of the optimization process described above. First, the feasible space is so remote, at least relative to the design space, that the optimizer is only guaranteed to find viable designs if it is initiated in close proximity to this space. Alternatively, it is possible that objective space possesses a degree of non-linearity that makes it difficult for the optimizer to identify the feasible space. Finally, local minima may impede the optimizer's ability of consistently migrating toward feasible regions. The challenge with these observations is that it is also quite possible, even likely, that no single factor is the cause of this complication. Rather, it is quite plausible that some combination of the three are driving the validation experiment to failure.

The process of designing experiments to test these hypothesis and then modify the optimization scheme based on the results would prove to be exceptionally complicated and time consuming. Fortunately, this approach is not required. The previous section already established that the NSGA-II based trade space algorithm is able to identify feasible regions of the space quite efficiently. This observation informs a new hypothesis - the existing trade space exploration process could be adapted to identify a particular point of interest as opposed to the entire frontier. Such an approach

www.manaraa.com

would likely be well suited to the task, given that meta-heuristic algorithms are well regarded for dealing with non-linear spaces and local optima. The challenged to be addressed is how this modification should be formulated.

There are two dimensions to the optimization problem: objectives and constraints. The OEC developed in the previous section provides the means to define the objective, but it is not immediately clear how constraints should be formulated. In a typical GA construct, constraints are often dealt with through a penalty function. As the name implies, this approach adds a "penalty" to a member of the population's OEC, where the magnitude of the penalty is proportional to the severity of the constraint violation. An important aspect of implementing this method is to calibrate this measure of proportionality. If the impact is too high, then the optimizer would be likely to either prefer points on the interior of the frontier, or high cost, high performance options, since both are unlikely to violate constraints. If, on the other hand, the penalty is not sufficiently severe, then the optimizer will accept points that violate constraints to some degree. The particular complication in this formulation is that each iteration of the MCS effectively creates a new frontier. As such, carefully calibrated parameters for one scenario may be insufficient for another. Another method is therefore likely to be required.

A potential solution could lie in the basic formulation of the NSGA-II algorithm. Recall that in this unique variation of the GA, members of the population that satisfy constraints are given priority over those that do not. This is the primary factor governing which members move onto the next generation. The objective function then serves as secondary consideration to determine which of the points with moderate performance should be retained. The specifics of the NSGA-II algorithm in implementing these concepts cannot be directly applied, but the basic concept is not hindered by the challenges which make the penalty function approach dubious.

Based on these observations, the following approach is proposed. When evaluating

المنارة للاستشارات
www.manaraa.com

a population of points, determine which members satisfy the constraint. If the number points meeting this criterion exceed the population size, then dismiss all points failing to meet constraints. Next, order the remaining members of the population according to the value of their objective function and iteratively remove the worst performing members until the correct population size is reached. If, on the other hand, there are not enough feasible members of the population to fill out the next generation, then apply the same concept in reverse. Specifically, order members of the population that failed the constraints according to their objective function value, and use the top performers in this category to fill out the ranks of the next generation. This approach would be consistent with the basic concepts of the NSGA-II algorithm, but the intent would be to drive the population to a specific point, rather than distribute those points along the frontier.

The technique just described was applied to the same baseline validation scenario used to vet the original continuous optimization scheme. Results from the first generation selection process are depicted in Figure 110 to visualize the implementation method for the proposed approach. There are a total of 25 points in this population (blue points) satisfying constraints, and 75 points violating constraints (red points). Since the population size is 50, the 25 points satisfying constraints immediately move to the next generation. Remaining points are evaluated with the OEC developed in the previous section, and the 25 points with the best score are allowed to proceed. The threshold value of the objective function defining these points is presented by the blue dividing line in Figure 111, and the points above this threshold (i.e. those moving on to the next generation) are highlighted.

Having fully described the implementation procedure of the new approach, the remaining consideration is to evaluate its performance on the baseline case. To that end, 1,000 iterations were run on the closed architecture frontier. Each iteration

المنارة للاستشارات
www.manaraa.com

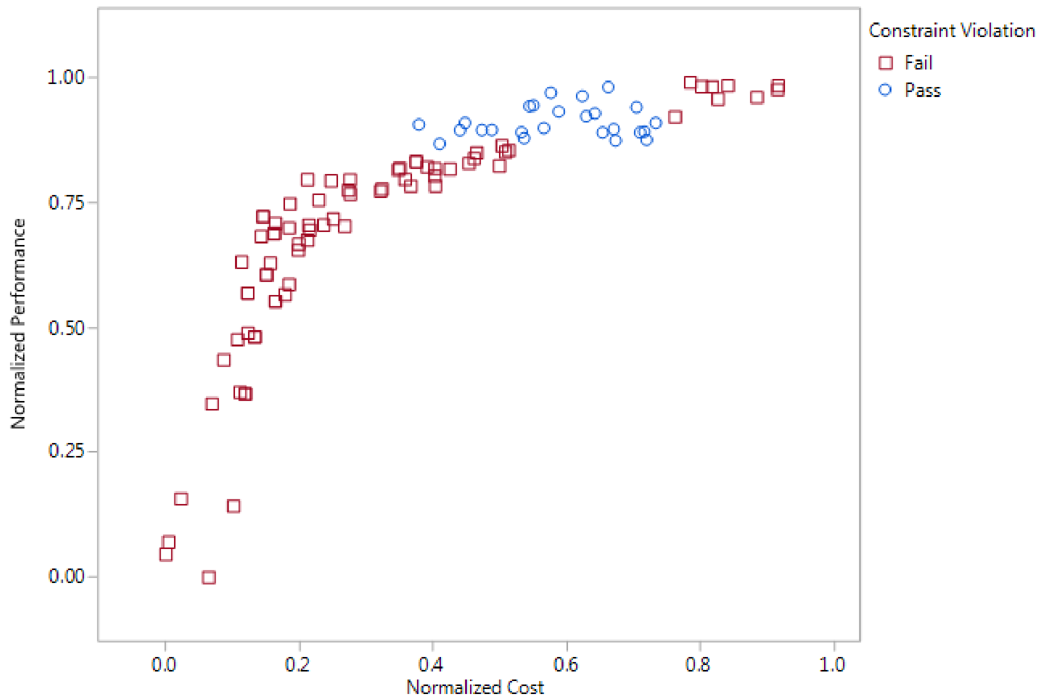


Figure 110: Identification of Feasible Designs in First Generation

is done on the expected scenario (i.e. no uncertainty) in order to isolate the variability created by applying a stochastic optimization. These results are presented in Figure 112 below, where the original frontier is represented in red, and repeated iterations of the optimizer for minimum cost and maximum cost-effectiveness are depicted in green and blue respectively.

In terms of finding the minimum cost, it is quite clear that the new optimizer performs extremely well, as results are tightly clustered on the frontier and in close proximity to the desired point. Results for finding the maximum cost-effectiveness point are less encouraging. The optimizer consistently finds the frontier, but the points are distributed across a much larger region than desired. Figure 113 demonstrates why this is likely to be the case. Here, iso-probability contours of results are depicted in blue, and are bounded by two constant OEC value functions in red. The value of the upper boundary in this figure is 0.362 and the value of the lower boundary is 0.305. This indicates that the optimizer is consistently converging to a similar

www.manaraa.com

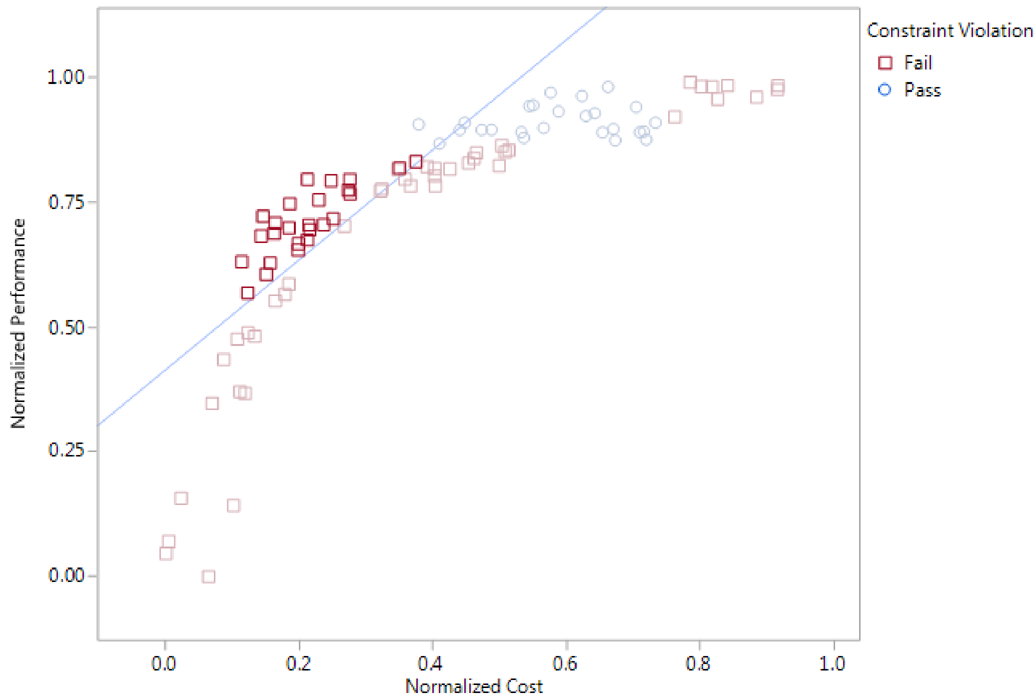


Figure 111: Down Selection of Non-Feasible Designs in First Generation

value, but the value can vary by 15% depending on the evolutionary path taken. This is not an unreasonable result, but the shallowness of the frontier in this region causes this moderate variation to spread across a wide area.

In conclusion, the proposed optimization scheme is vastly superior to the original formulation, but the added stochastic variation remains problematic. The impact of this variation is likely dependent on the shape of the frontier in proximity to the design point being investigated, which implies that its significance will vary from problem to problem. This short coming should be addressed in future work, but it will be assumed to be sufficient for the remainder of this analysis.

المنارة للاستشارات
www.manaraa.com

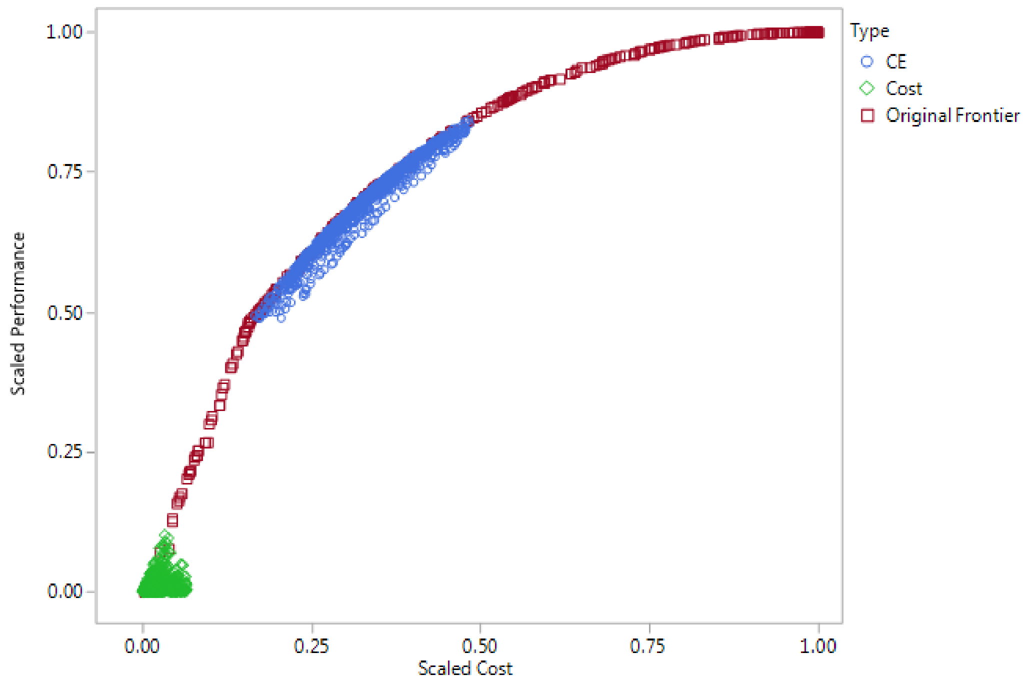


Figure 112: Variability Introduced By Stochastic Optimization

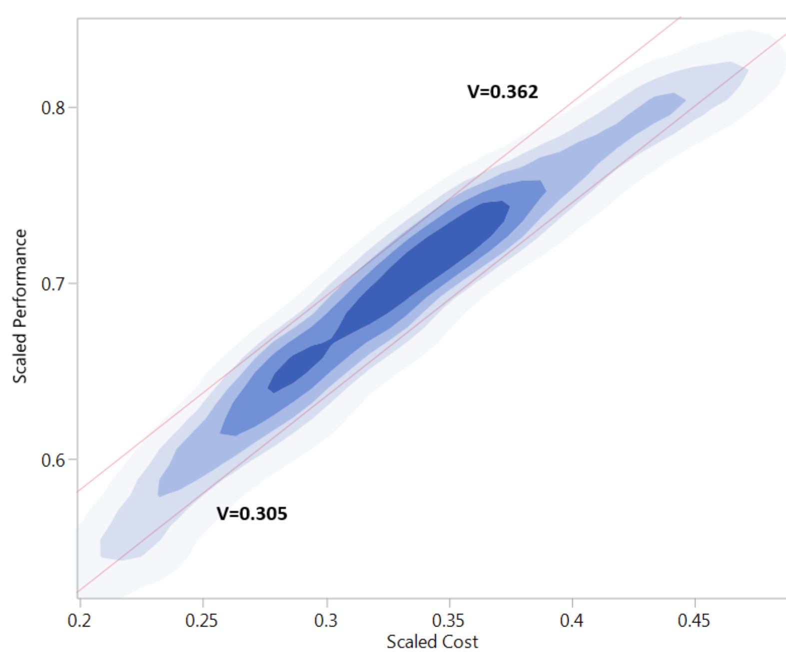


Figure 113: Change in Objective Function Value for Cost-Effectiveness Search

المنارة للاستشارات
www.manaraa.com

8.6.4 Monte Carlo Results

With the new optimizer defined and validated, the remaining step is to apply it to a full MCS and evaluate the results for a single upgrade. Figure 114 presents these results as a contour plot of 2,500 MCS iterations, where the same convention from the previous section is retained: green regions indicate the density of points for the minimum cost strategy, and blue represents the density of points for the maximum cost-effectiveness strategy. Figure 115 presents the same results as a scatter plot.

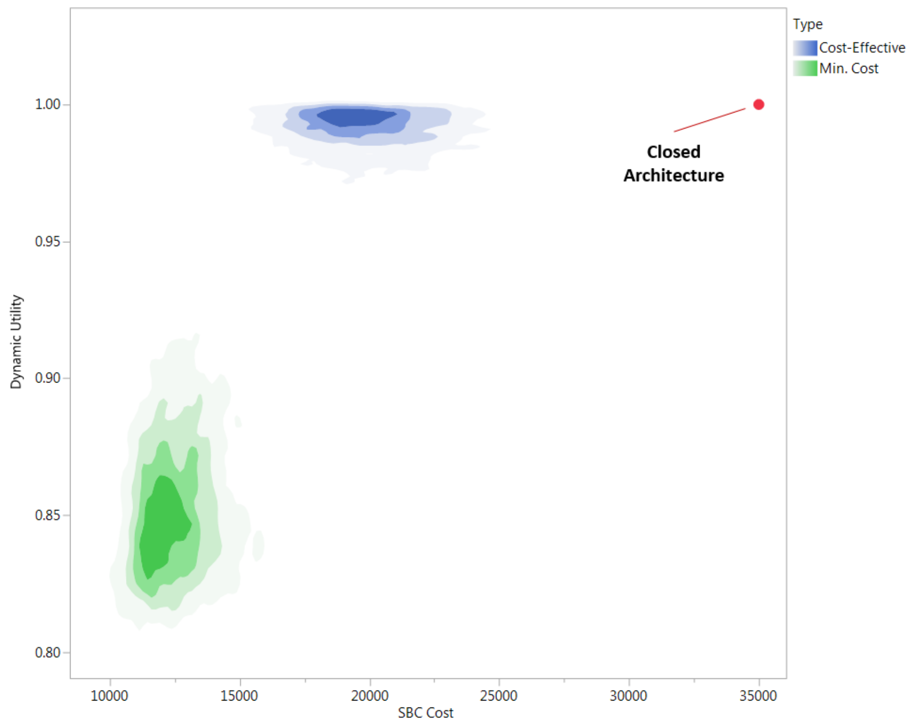


Figure 114: Four Level Contour Representation of MCS Results

There are several observations to be drawn from Figures 114 and 115. First, both strategies appear to have a similar measure of variability in the distribution of their respective MCS results. They key distinction is that the major axis of variability for the minimum cost approach is the performance dimension, while the opposite is true of the maximum cost-effectiveness strategy. This visual trend is confirmed and quantified in Table 15, which presents the mean and standard deviation for the

المنارة للاستشارات
www.manaraa.com

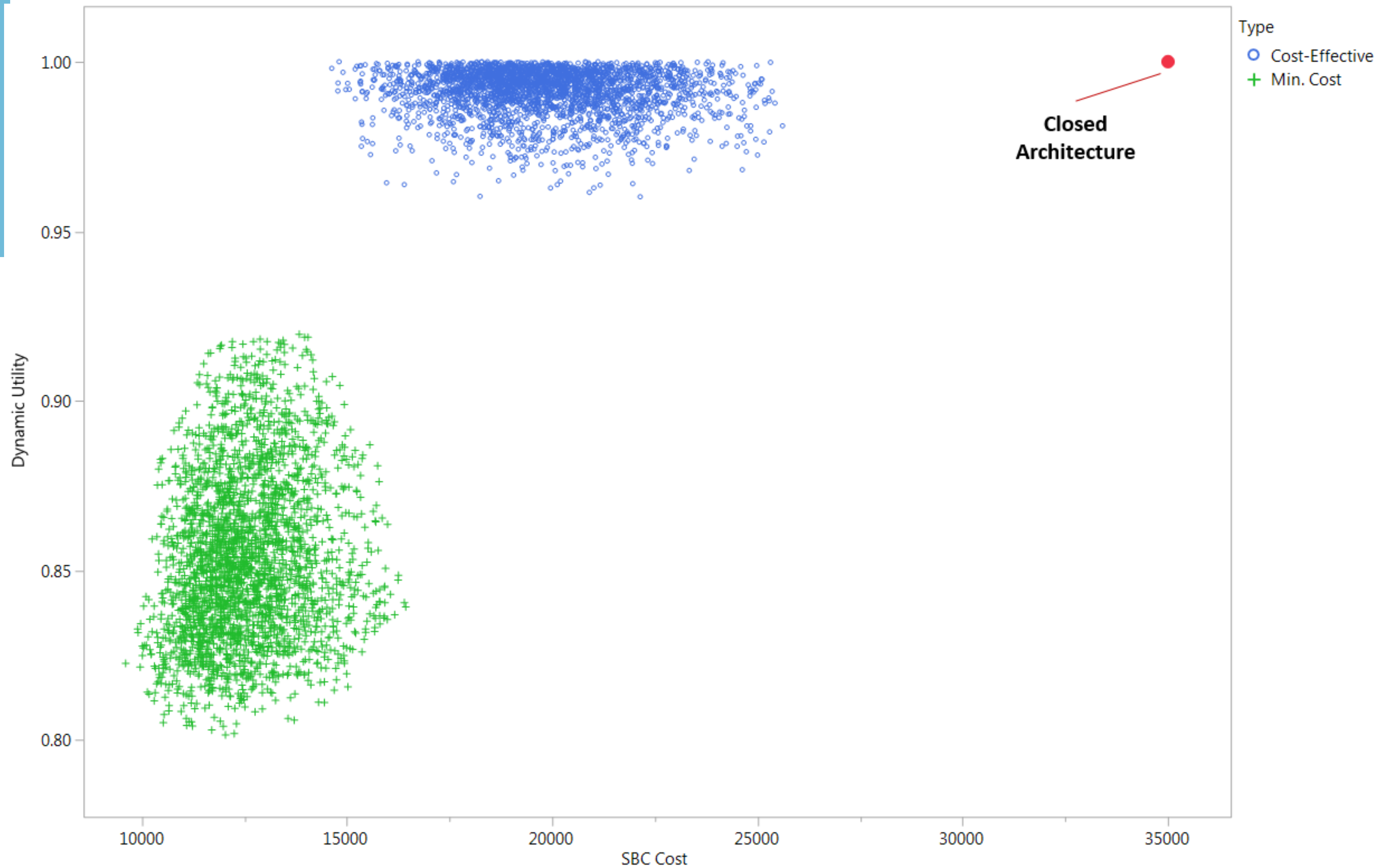


Figure 115: Scatter Plot Representation of MCS Results

www.manaraa.com

marginal distributions (i.e. performance and cost) associated with each strategy. These results indicate that the minimum cost approach would allow the DoD to, on average, reduce unit costs by 37%, but this comes with at the expense of a 16% reduction in performance. In addition, the ratio of standard deviations indicate that a minimum cost approach would have a four fold increase performance uncertainty compared to the cost-effective strategy, but 50% less uncertainty in cost. This is an interesting result. The general trend seems to be that emphasizing cost as the key metric increases uncertainty in the actual performance that will be achieved. On the other hand, emphasizing cost-effectiveness provides much greater stability in system performance, but substantially increases uncertainty in the cost of maintaining that performance.

Table 15: Marginal Distribution Statistics for Alternative TRP Strategies

	Minimum Cost		**Cost-Effective**	
	Mean	Std. Dev.	Mean	Std. Dev.
Performance	0.855	0.025	0.991	0.006
Cost	12,555	1,200	19,822	1,971

Next, it is interesting to note the relative position of the closed architecture design in both figures. This design is generated by requiring the system to provide a 99% confidence of satisfying performance requirements, given by the upper confidence boundary in Figure 76, without infusing technology at a later date. The benefit of this approach lies in the fact that there is a negligible degree of uncertainty in results; cost are know with absolute certainty, and performance targets are, by definition, completely satisfied in 99% of cases. However, this approach is substantially more expensive that either upgradeable strategy. Further, the difference in performance between the closed architecture and the cost-effective TRP is, on average, less than 1%.

This is a very important observation. The general argument opposing OSA design

المنارة للاستشارات
www.manaraa.com

is that there is too much uncertainty in the outcome of future events to justify the cost and risk of implementing an open design. In this scenario, it is true that there is a significant degree of uncertainty in the cost of associated with pursuing an open approach. Yet, even in the worst case scenario, this cost is far less than the cost associated with developing an extremely high performance system in the present. Many authors argue this point in the OSA design literature, but this is the first instance in which the argument can be quantified and documented in a rigorous fashion.

Finally, the previous section observed that the stochastic optimizer was likely to add significant variability to MCS iterations evaluating the cost-effective approach. It was therefore assumed that the cost-effectiveness results would have greater variability compared to their minimum cost counterpart, and that this trend would extend to both axes. This hypothesis proved to be true in the cost dimension, but not, as implied by the original study, in the performance dimension. Given this observation, it is possible that the variability introduced by the optimizer is far less than the that generated by uncertainty in the MCS scenario. However, this hypothesis requires further study to support.

المنارة للاستشارات
www.manaraa.com

CHAPTER IX

CONCLUSION

9.1 Review

The purpose of this research was to develop a methodology to address one of the unfortunate trends that has plagued military acquisitions for generations, namely the exponential growth in costs associated with moderate improvements in capabilities. There are numerous factors contributing to this trend, and it is quite unlikely that no "silver bullet" exists to simultaneously address these factors. However, are there concepts within this field that could potentially be addressed through an improved Systems Engineering framework: Gold Plated requirements, Vendor Lock, and the need for custom components. The open architecture approach was developed over two decades ago to address these specific factors. Embracing this methodology would reduce the cost and complexity of upgrading systems, thereby allowing system requirements to be derived from an earlier stage in the system life cycle where uncertainty is reduced. In addition, mandating the use of widely accepted commercial standards and open interfaces not only maximizes the use of existing COTS products, it also increases the pool of vendors available to bid on sustainment contracts throughout the system life cycle. Combined, these properties would effectively break the self fulfilling prophecy that invariably leads the military to develop expensive, highly integrated systems that contribute heavily toward the trend of exponential cost growth.

While the open architecture approach is conceptually simple, it has proven to be exceptionally difficult to implement in practice. The broader acquisitions community has documented several barriers to entry that prevent the idealized open system from becoming a realistic scenario. In essence, these barriers relate to a series of simple

www.manaraa.com

observation: (1) imposing open architecture constraints increase cost and risk during the development process; (2) benefits provided by this type of system accrue in the future and are subject to considerable uncertainty; (3) there is no consistent method to determine if the uncertain value provided in the future is worth the cost and risk assumed in the present. Recognizing these challenges, several organizations have developed decision support methods to help facilitate this determination. However, many in the acquisitions community find these methods to be of dubious values, since these they rely almost entirely on the opinions of SME's, reduce numerous factors into simple qualitative metrics, and provide no mechanism to incorporate uncertainty. These observations provided the foundational motivation for this work, which is expressed in the follow objective statement:

> **Research Objective: Allow for more informed tradeoffs between open and closed system system architectures by developing a methodology to measure the expected costs, benefits, and risks associated with upgradeable systems. The methodology should provide a means to incorporate the likely evolution of commercial technologies and system requirements, uncertainties inherent to forecasts of future events, and the managerial flexibility to alter decisions as a hedge against uncertainty.**

Subsequent analysis decomposed this research objective into three research questions that must necessarily be resolved to successfully satisfy the overall objective. Chapter Four considered bias mechanisms that, if not properly addressed early in the development process, sow the seeds of failure in the earliest stages of design. Though there were many possible bias mechanisms, this work focused on those mechanisms that engineers can, and therefore should, address with rigorous SE methods: proper definition of system requirements, and development of a comprehensive modular partitioning scheme.

المنارة للاستشارات
www.manaraa.com

To properly formulate requirements, it was determined that the current practice in existing open architecture design methods of documenting the evolution in system requirements overtime as a CRM is necessary, but not sufficient. This lack of sufficiency stems from the fact that there is substantial uncertainty inherent to these forecasts, and current methods of modeling this uncertainty are too qualitative and subjective for decision-makers to have confidence in their results. Several quantitative approaches were identified in the existing literature with the potential to provided greater transparency and objectivity in this analysis. The review concluded that a parsimonious approach to elicit and model this uncertainty is to engage SME's in an iterative linguistic bounding and fit process. This method provides a consistent, objective, and repeatable means to map expert opinion into a closed form mathematical representation, which can then be directly integrated into probabilistic models.

With respect to modularity, any portion of the system intended to be upgraded over time must be partitioned from the greater system architecture as a self-contained module. It was observed that the broader field of modular design offers several methods to measure the extent to which modularity is present in a given partitioning scheme. However, there is no academic consensus as to how one would determine which method is the most appropriate, nor is there an accepted threshold to determine if a given value is acceptable. In short, there is no panacea to resolve this challenge. The proposed methodology therefore advocates that an acceptable approach to manage this complexity can be formulated as a three-stage process. First, the Design Rules technique developed by Baldwin and Clark is leveraged to decompose the greater system architecture into independent sub-systems. Second, functional heuristics are applied to individual sub-systems in order to identify an elementary set of basis modules. Finally, the functional resources provided by these elementary units are compared to the requirements expressed in the Capability Road Map in order to identify which components warrant further consideration. This approach is

المنارة للاستشارات
www.manaraa.com

advantageous in that it provides a consistent process, is applicable to a wide variety of systems, and reasonably accommodates designs with substantial complexity.

The second research question was considered in Chapter Five, which began with the recognition that acquisition decision-makers will not accept that an OSA will efficiently satisfy evolving requirements on faith alone. Rather, open systems must undergo the same cost-effectiveness scrutiny to which closed system architectures are subjected. This requires, among other things, a TRP to be presented alongside the system concept to describe when components will be upgraded, what the resulting effectiveness will be relative to projected thresholds, and what TLCC should be expected given that these upgrades occur. While the timing of upgrades could be arbitrary, the increase in system effectiveness and its corresponding cost are not. Developing a TRP therefore requires a method to forecast the evolving properties of maturing technologies.

Several competing forecasting methods were identified in the literature, but only a select few accommodate technologies defined by multiple measures of performance. Within this category, it was ultimately determined that the MDGM approach provided the most efficient means of approximating the state of the art at a future point in time. This approximation is mathematically incorporated into the analysis as a constraint on the values of design variables that are accessible at the time in question. As time passes, the constraint is relaxed and a larger portion of the theoretical design space becomes accessible, which provides a means through which the performance of system can be improved. In addition, a statistical CER is used to estimate the cost of technology as a function of its design variables. In this way, an analyst can construct a realistic TRP by selecting the timing of a given infusion and the desired design variables. Further, since an improvement in design variables would propagate through M&S to generate an increase in response variables, it now becomes possible to visualize the impact of a TRP as a time history of system performance.

المنارة للاستشارات
www.manaraa.com

While this representation is advantageous in conveying the impact of complex design decisions, it is not a useful metric for decision making purposes. Generating such a metric requires a means to determine when one TRP is preferable to another in order to screen out inefficient concepts. This work advocates the use of Dynamic Value Theory for this purpose. Dynamic Value Theory is governed by two fundamental assertions: the instantaneous value of a system is the difference between what the stakeholders desire and what the system can deliver over time; and stakeholders base their decisions on life cycle value, which is defined as the integration of instantaneous value over the period of interest. These assertions can be related to the refresh planning problem by recognizing that the stakeholder's desires are contained in the CRM, and the system's capabilities are contained in the TRP. Several dynamic valuation concepts are formulated under this construct, each of which serves to reduce the time history of response variables to an equivalent performance metric.

The ability to automatically generate and compare alternative TRP's affords provides the means through which the trade space of alternative TRP's can be defined. The approach used in this work is a variation of a popular bi-level search technique applied in PPD literature. In this formulation, the top level passes a set of timing assumptions to a lower level optimizer implementing the NSGA-II template. Aggregating efficient points found across all timing assumptions then provides the trade associated with the open architecture system as a whole. Decision-makers can then evaluate alternatives in the same way in which static system architectures would be evaluated.

Chapter Six address the final research question, which considers the balance between uncertainty inherent to forecasts and the flexibility to adapt decisions in response to new information. This concept is not well addressed in traditional design literature, but there is a strong analogy between the flexibility provided by upgradeable systems and flexibility provided by financial options. However, a review of these

المنارة للاستشارات
www.manaraa.com

methods indicates that they are unsuitable in their native form, which is predicated on a cost-benefit valuation scheme. Yet, it was observed that all methods paralleling the open architecture design problem embrace the common heuristic of pricing by simulation. This heuristic was adapted to the problem under consideration, which provides a mechanism to alter decisions in response to alternative scenarios. This mechanism was then embedded in a classical MCS to consider the true impact of uncertainty on open design concepts.

The proposed methodology was then evaluated against a real world example problem in which an ISR platform was expected to receive periodic sensors upgrades. Results were broadly encouraging, as the proposed methodology was able to model the problem without modification. There were, however, two instances in which the initial approach failed to consider important factors. First, it was determined that correlation among component performance metrics can introduce complications when attempting to fit the relevant models to real world data. This problem appears to be well managed by replacing design variables with their principle components during the fit process, and this technique was added to the baseline approach. Second, the original MCS framework advocated using a continuous optimizer to manage the process of altering decisions across different scenarios. Experimental evidence indicates that this approach is ill-advised. An alternative, stochastic optimization framework was developed to circumvent the short comings of the deterministic method, which demonstrates adequate performance.

Once these challenges were resolved, the proposed methodology performed as intended. In addition, the results support many of the qualitative arguments made by proponents of open system design, which lends considerable support to central thesis of this work. Finally, when comparing to the proposed methodology to the KOSS and Risk Assessment methods presently in use by the acquisitions community, it is quite clear that the formulation in this work is a substantial improvement with respect to

المنارة للاستشارات
www.manaraa.com

the implementation barriers noted in Chapters One and Two. As such, the proposed methodology should be considered as a best practice for future open system design work.

9.2 Future Work

A typical dissertation tends to have a very narrow focus on specific attributes of a well posed problem. Reflecting on the wide range of topics that were addressed throughout the development of the proposed methodology, it is clear that this work does not fit that mold. The wide scope of literature considered in the earlier chapters (i.e. modular design, technology forecasting, dynamic value, meta-heuristic optimization, product platform design, etc.) was necessary because there is simply no existing framework to aggregate the various factors into a well posed problem. Thus, the principle contribution of this work lies in the formulation of a well posed problem that can serve as the foundation for future work. Consequently, there is room for substantial improvement in each of the modules that constitute the methodology, and the hope is that future work will continue to refine the approach outlined in this work.

With that in mind, a significant area of improvement for this work relates to the manner in which analysts define the problem to be addressed through future technology infusions. A key element in performing a realistic analysis of alternatives is ensuring that the scenarios are realistic, and that a broad range of alternative scenarios are identified. This work posits the existence of a Functional SME to provide all of the requisite information necessary to character the scenario, but this assumption is likely to be imperfect. A better approach would be to derive the necessary information directly from the strategic road maps that are commonly generated by DoD organizations. There are two apparent barriers that prevent this from occurring. First, strategic road maps are often geared toward resolving capability gaps

المنارة للاستشارات
www.manaraa.com

that require revolutionary advances in technology to resolve. Yet, it is unlikely that an existing system, regardless of whether its architecture is open or closed, will be able to integrate revolutionary technology without substantial, and highly expensive, modifications. This approach would therefore undermine the utility of an open architecture. Second, strategic road maps are often intentionally vague. This ensures that all available development paths are considered, but it prevents any realistic planning that would facilitate integrating new technology. With this in mind, an important area of future work should focus on developing a bridge between strategic road maps and system level CRM's.

As previously mentioned, another area of improvement in the proposed methodology is the stochastic optimization framework used to adapt decisions in the sensitivity analysis. A common challenge in probabilistic analysis is the inability to identify how much variability in the final output is caused by a particular variable. This is true of the work performed in Chapter Eight as well. Specifically, the intent of the sensitivity analysis is to isolate the variability in results caused by uncertainty in the co-evolution of technology and requirements from uncertainty in the decision process. The imperfect nature of the stochastic optimizer impedes this objective.

Another important observation from Chapter Eight is the need to implement a realistic sustainment model to map components costs over time to a TLCC. This is not a trivial endeavor. Unlike military specific components, the DoD has no control over commercial production, and the shelf life of these components is invariably much shorter than the life cycle of the military platform they would support. NRE costs for design and recertification are incurred, as with the implementation of a TRP, whenever COTS obsolescence requires a new component to be integrated into the system. Controlling these costs requires a well formulated obsolescence mitigation strategy to determine when these recertifications occur, and the number of components required

المنارة للاستشارات
www.manaraa.com

to sustain the fleet between transitions. Moreover, the cost model used in the experimental scenario assumes that NRE costs are negligible, and that the purchase rate is constant over the system life cycle. While this is sufficient to determine the main effects of cost trends, it is clear that TRP and the obsolescence mitigation strategy cannot be decoupled. Further work is therefore required to determine how such a plan could be formulated, and how it could be integrated with the TRP methodology developed in this work.

www.manaraa.com

APPENDIX A

OAAT QUESTIONNAIRE: PROGRAMMATIC

1.1	To what extent is OSA incorporated into the program's acquisition planning?
1.2	To what extent did the program plan for its implementation of OSA?
1.3	To what extent is the program's OSA implementation based on systems engineering principles and resources?
1.4	To what extent are responsibilities assigned for implementing OSA?
1.5	To what extent is the program staff trained on, or have relevant experience in OSA concepts and implementation?
1.6	To what extent does the program's configuration management process encompass changes to key interfaces and corresponding standards?
1.7	To what extent have program requirements been analyzed, and refined as needed, to ensure that design-specific solutions are not imposed?
1.8	To what extent do the system level functional and performance specifications permit an open system design?
1.9	To what extent are modular, open system considerations included as part of alternative design analyses?
1.10	To what extent are mechanisms established to migrate key interfaces that are proprietary or closed to key interfaces that are open?
1.11	To what extent are OSA reflected in the program's performance measures?
2.1	To what extent does the program have policies and processes that control adding specifications, options, or extensions that limit the use of widely-supported or openly available interface standards?

المنارة للاستشارات
www.manaraa.com

2.2	To what extent are non-mission unique capabilities supplied using either components reused from other programs or available from the commercial market?
2.3	To what extent have the proprietary or unique non-commercial elements been limited or well defined such that they do not hinder other developers from interfacing or developing any part of the system?
2.5	To what extent is the program complying with the Joint Capability Integration and Development System (JCIDS)?
2.6	To what degree is the program complying with the Interoperability and Supportability requirements for national security systems in references like CJCS 6212.01C and DoDD 4630.5 and DoDI 4630.8?
2.7	To what extent does the program plan directive documentation and funding enable orderly migration of proprietary or program unique system modules to open system alternatives when capabilities are upgraded?
2.8	To what extent is the program free of system components that have proprietary characteristics, restrictive licensing or prohibitive cost that could limit or preclude the reuse of the components in other Navy systems or the competitive selection or re-assignment of those components to other vendors?
2.9	To what extent has the Prime System Integrator established processes that facilitate flexibility of task assignment, competition of individual tasks, or re-competition of tasks?
2.10	To what extent has the program established and maintained a repeatable, non-restrictive process that discloses in-process design documentation and software tools information directly to third party developers?
2.11	To what extent is design documentation disclosed to interested parties from the beginning of the development effort?

المنارة للاستشارات
www.manaraa.com

2.12	To what extent does the Program documentation stress the use of widely-accepted and supported standards, such as those maintained by recognized organizations (e.g. IEEE), to define both internal and external interfaces?
2.13	Does the program plan and directive documentation include a data management strategy ensuring that when the Government exercises its intellectual property rights to obtain any developmental artifacts for anything it paid to develop with either complete or partial funding the Contractor can at most charge a nominal fee covering the marginal cost to the effort to provide that documentation?
2.15	To what extent does the program plan include lifecycle support and funding for open architecture elements?
2.16	To what extent has the program worked with the applicable Tech Warrant holder or equivalent authority to develop open architecture specific metrics as part of its program management processes and reviews?
2.17	To what extent do the program's software selection criteria require that, other things being equal, priority be given to software components/modules/systems that have the least restrictive rights associated with them?
2.18	To what extent does the program specify that system components have well-defined interfaces, information exchange standards, functional requirements and specifications?
2.19	To what extent does the program use OSA-specific language or contractual provisions in its acquisition and development documentation?
2.20	To what extent does the program's documentation provide for cost-effective incremental upgrades without dependencies on a single source or the need to redesign large portions of the system?

المنارة للاستشارات
www.manaraa.com

2.21	To what extent has the program organization implemented a training program to educate their workforce on OSA-related policy and concepts?
2.22	To what extent has the program used incentives to promote modular designs, commonality and component reuse?
2.23	To what extent does the program's configuration management process use integrated teams to identify how individual changes impact the system's interfaces and information exchange standards?
2.24	To what extent are multiple third parties directly contracted to develop components of the system, giving the government the flexibility to compete or reassign component development?
2.25	To what extent do the program's market research and selection processes use criteria that favor commercial, common enterprise wide, or generally accepted interface and information exchange standards?
2.26	Does the program's acquisition strategy, contract language and funding profile facilitate subsequent assignment of major tasks and program roles to alternate providers at predetermined intervals?
2.27	To what extent are market research, community of interest teams, peer review groups, or alternative forums used to access and select among available capability improvement options?
2.28	To what extent does the program develop POM issue papers or other business planning documents to address OSA business and technical issues?
2.29	To what extent does the program reuse components from other government programs?
2.30	To what extent can the program accomodate software tools or other components from sources other than the prime system integrator or existing vendors without requiring significant modifications?

المنارة للاستشارات
www.manaraa.com

APPENDIX B

OAAT QUESTIONNAIRE: TECHNICAL

3.1	To what extent is the system's architecture based on related industry or other standard reference models and architectural frameworks?
3.2	To what extent is an architectural description language used to define system modules and interfaces?
3.3	To what extent does the system's architecture exhibit modular design characteristics?
3.4	To what extent is the system's architecture capable of adapting to evolving requirements and leveraging new technologies?
3.5	To what extent has the criteria for designating key interfaces been established?
3.6	To what extent has the program designated key interfaces?
3.7	To what extent has the program assessed the feasibility of using open standards for key interfaces?
3.8	To what extent have standards selection criteria been established that give preference to open interface standards?
3.9	To what extent are open standards selected for key interfaces?
3.10	To what extent are validation and verification mechanisms established to assure that system components and selected commercial products conform to the selected interface standards?
3.11	To what extent do system components and selected commercial products conform to standards selected for system interfaces?
3.12	To what extent do system components and selected commercial products avoid utilization of vendor-unique extensions to interface standards?

المنارة للاستشارات
www.manaraa.com

3.13	To what extent can system components be substituted with similar components for competitive sources?
4.1	The unit of assessment predominantly complies with what type of interoperability standards?
4.3	What is the scope of the data model that the unit of assessment uses to support interoperability with other systems?
4.4	What is the scope of interoperability of the unit of assessment?
4.5	To what extent does the unit of assessment use mechanisms to discover and invoke services?
4.6	To what extent does the unit of assessment support mechanisms for service discovery and invocation?
5.1	What architectural characteristics address obsolescence and provide for timely technology refresh, fixes and upgrades?
5.2	Do the unit of assessment's technical artifacts provide sufficient detail and scope for maintenance?
6.1	Does the program follow a well-defined Systems Engineering process for implementing capability extension?
6.2	Will the technical infrastructure accommodate extensibility of the unit of assessment's functionality?
6.3	What is the scope of testing needed after new components are added to the unit of assessment?
7.1	To what extent are the components of the unit of assessment implemented and independently deployable as packages?
7.2	To what extent can the functional capabilities of the unit of assessment be re-combined or re-arranged to support a modified process/workflow/mission?
8.1	What reuse strategy is used within the unit of assessment?

المنارة للاستشارات
www.manaraa.com

8.2	What is the scope of the set of processes used to identify and evaluate reuse candidates for incorporation into the unit of assessment?
8.3	Which approach best describes the operational run-time infrastructure supporting the unit of assessment?
8.4	Have the commonalities and variations of the unit of assessment been specified to facilitate reuse congruent with a broader software product line?

المنارة للاستشارات
www.manaraa.com

APPENDIX C

CDF OF SCENARIO PARAMETERS

The figures presented below characterize the uncertainty present in modeling assumptions for the experimental proof of concept in Chapter Eight.

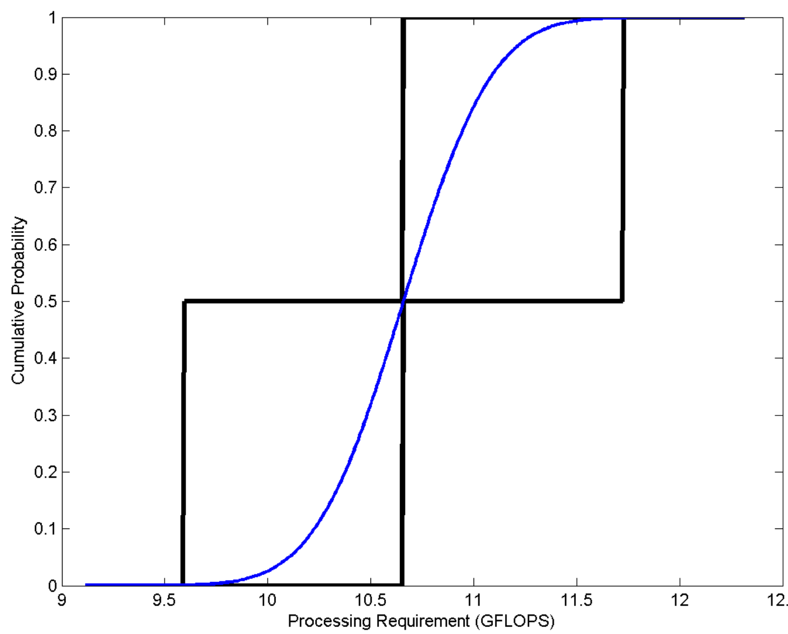


Figure 116: CDF of Requirements Assumptions- First Capability Increment

www.manaraa.com

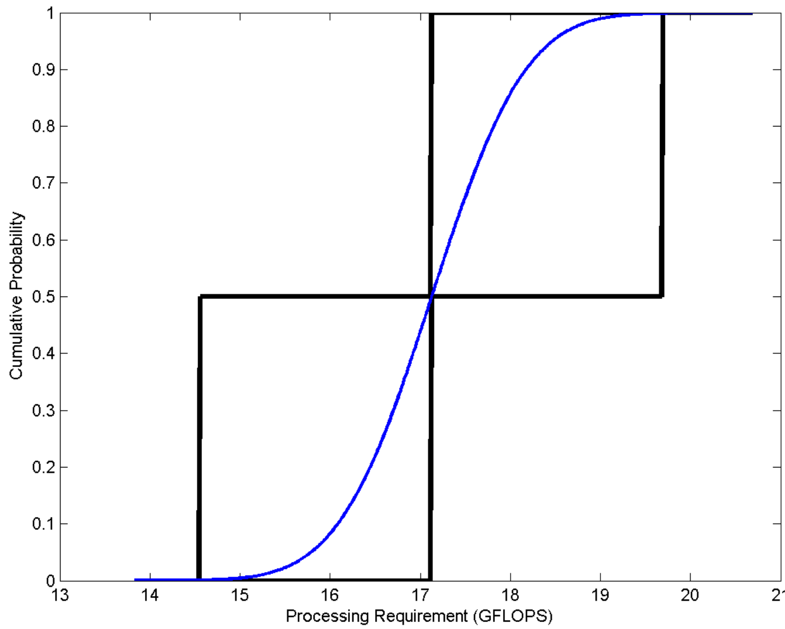


Figure 117: CDF of Requirements Assumptions - Second Capability Increment

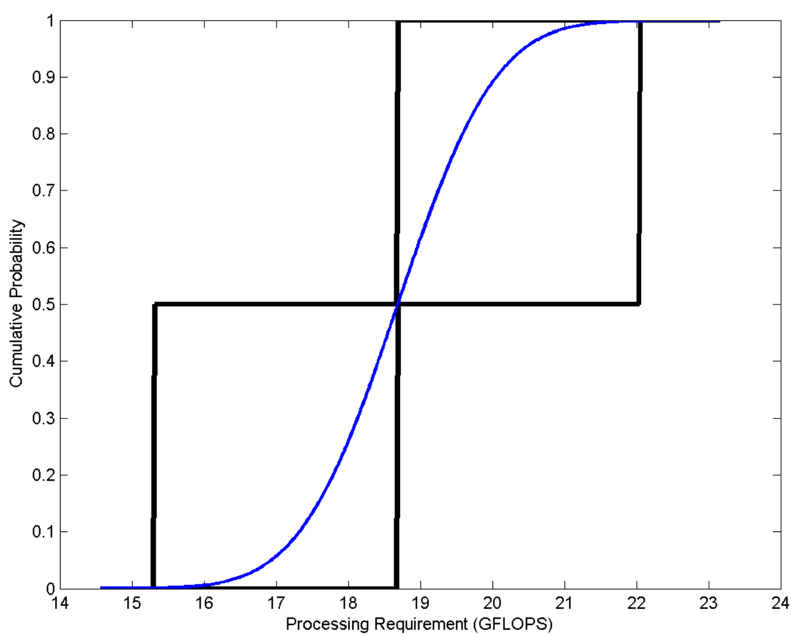


Figure 118: CDF of Requirements Assumptions - Third Capability Increment

www.manaraa.com

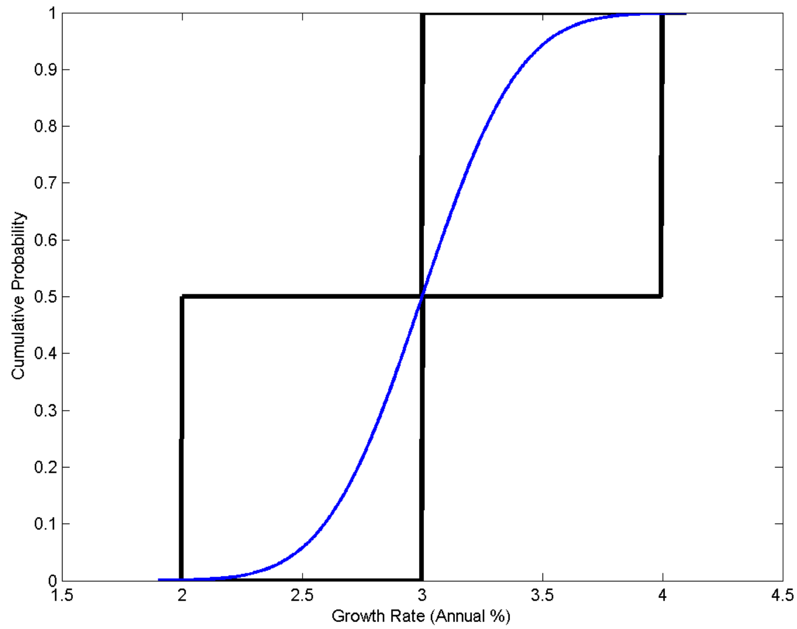


Figure 119: CDF of Requirements Assumptions - Annual Growth Rate

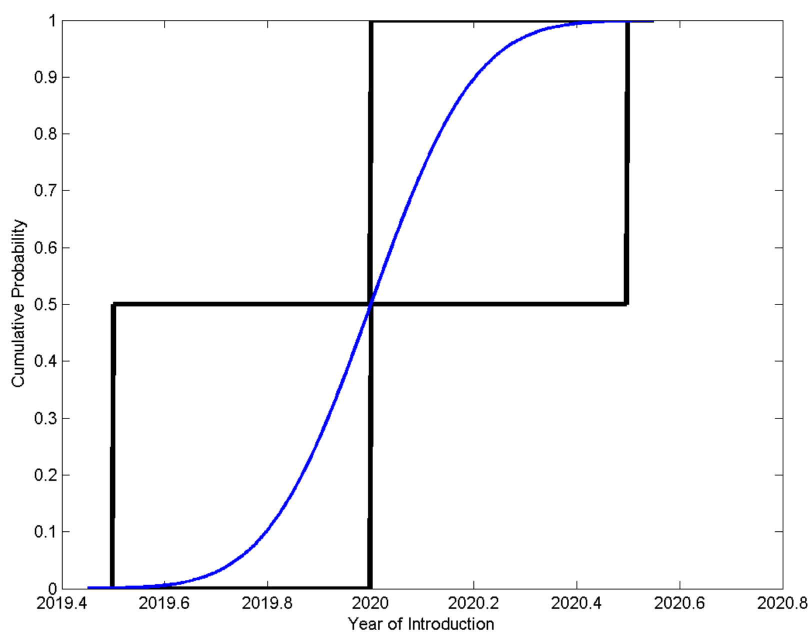


Figure 120: CDF of Timing Assumptions - First Capability Increment

www.manaraa.com

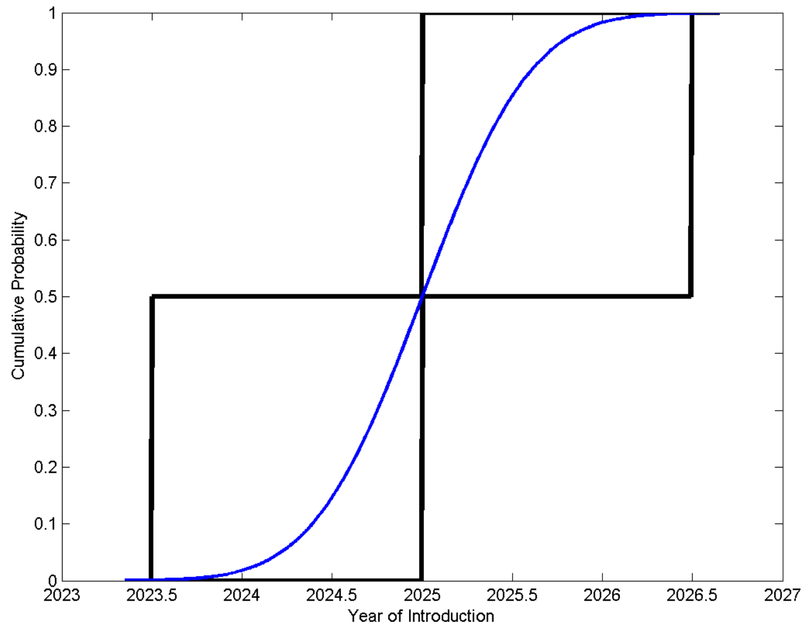


Figure 121: CDF of Timing Assumptions - Second Capability Increment

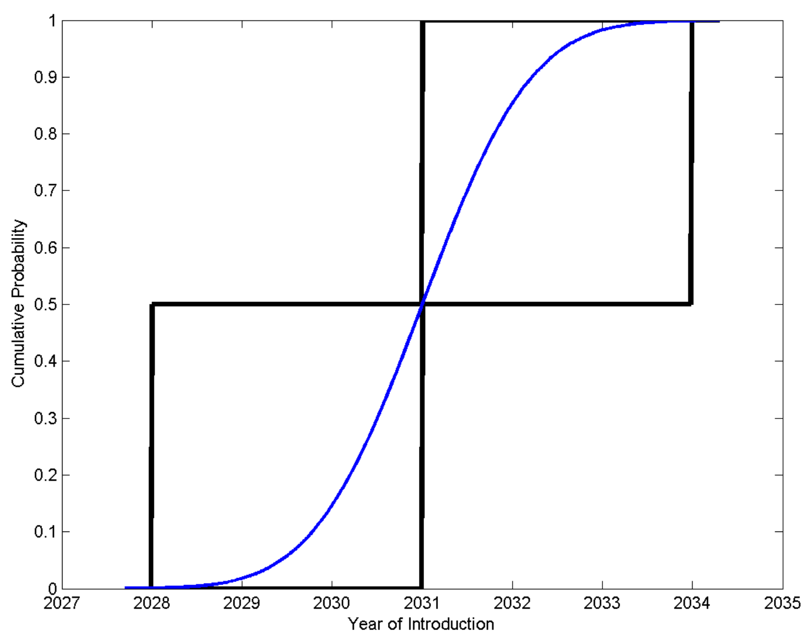


Figure 122: CDF of Timing Assumptions - Third Capability Increment

www.manaraa.com

REFERENCES

[1] Agarwal, H., *Reliability Based Design Optimization: Formulation and Methodologies.* PhD thesis, University of Notre Dame, December 2004.

[2] Akundi, S., Simpson, T., and Reed, P., "Multi-objective design optimization for product platform and product family design using genetic algorithms," in *ASME Design Engineering Technical Conferences - Design Automation Conference*, no. DAC-84905 in DETC2005, (Long Beach, CA), ASME, 2005.

[3] Alexander, A. and Nelson, J., "Measuring technological change: Aircraft turbine engines," *Technological Forecasting and Social Change*, vol. 5, pp. 189–203, 1973.

[4] Allen, K. and Carlson-Skalak, S., "Defining product architecture during conceptual design," in *Proceedings of the 1998 ASME Design Engineering Technical Conference*, (Atlanta, GA), 1998.

[5] Anderson, L., "A simple approach to the pricing of bermudan swaptions in the multifactor libor market model," *Journal of Computational Finance*, vol. 3, no. 5, pp. 5–32, 2000.

[6] Anderson, T., Grosskopf, S., Far, R., and Song, X., "Further examination of moore's law with data envelopment analysis," *Technological Forecasting and Social Change*, vol. 69, pp. 465–477, 2002.

[7] Anderson, T., K.Hollingsworth, and Inman, L., "Assessing the rate of change in the enterprise data system market over time using dea," in *Technology Management in the Knowledge Era*, PICMET, (Portland, OR), pp. 384–390, 2001.

[8] Aounes, Y. and Chateauneuf, A., "Benchmark study of numerical methods for reliability-based design optimization," *Structural and Multidisciplinary Optimization*, vol. 41, pp. 277–294, 2010.

[9] Arena, M., "Why has the cost of navy ships risen? a macroscopic examination of the trends in u.s. naval ship costs over the past several decades," tech. rep., RAND Corporation, 2006.

[10] Augustine, N., *Augustines Laws.* American Institute of Aeronautics and Astronautics, sixth edition ed., 1997.

[11] Ayyub, B., *Elicitation of Expert Opinions for Uncertainty and Risk.* CRC Press, 2001.

المنارة للاستشارات
www.manaraa.com

[12] Azani, C., "A multi-criteria decision model for migrating legacy system architectures into open systems and systems-of-systems architectures," *Defense Acquisition University*, pp. 320–341, October 2009.

[13] Azani, C. and Khorramshahgol, R., "Management of technology using a modular open system strategy," in *Proceedings of the 12th European Concurrent Engineering Conference*, (Toulouse, France), pp. 5–12, 2005.

[14] Azani, C. and Khorramshahgol, R., "Modular open systems approach: An effective business strategy for building affordable and adatapable architectures," *Journal of Management Systems*, vol. 18, no. 1, pp. 66–76, 2006.

[15] Baldwin, C. and Clark, K., *Design Rules: The Pointswer of Modularity*, vol. 1. MIT Press, 2000.

[16] Bandte, O. and Mavris, D., "Multi-objective optimization using a joint probabilitstic technique," in *8th Symposium on Multidisciplinary Analysis and Optimization*, (Long Beach, CA), AIAA/NASA/USAF/ISSMO, September 2000.

[17] Barraquand, J. and Martineau, D., "Numerical valuation of high dimensional multi-variate american securities," *Journal of Financial and Quantitative Analysis*, vol. 30, no. 3, pp. 383–405, 1995.

[18] Bass, F., "A new product growth model for consumer durables," *Management Science*, vol. 15, pp. 215–227, 1969.

[19] Bean, E., Kerner, H., Patout, B., Davis, L., and Robertson, M., "Defense acquisition reform: Where do we go from here?," staff report, U.S. Senate: Committee on Homeland Security and Governmental Affairs, October 2014.

[20] Black, F. and Scholes, M., "The pricing of options and corporate liabilities," *The Journal of Political Economy*, vol. 81, pp. 637–654, May-June 1973.

[21] Blackman, A., "The rate of innovation in the commercial aircraft jet engine market," *Technological Forecasting and Social Change*, vol. 2, pp. 269–276, 1971.

[22] Blackman, A., "A mathematical model for trend forecasts," *Technological Forecasting and Social Change*, vol. 3, pp. 441–452, 1972.

[23] Bolten, J., Leonard, R., Arena, M., Younossi, O., and Sollinger, J., "Sources of Weapons System Cost Growth: Analysis of 35 Major Defense Acquisition Programs," Project Air Force MG-670-AF, RAND Corporation, Santa Monica, CA, 2008.

[24] Brandao, L., Warren, J., and Hahn, J., "Using binomial decision trees to solve real-option valuation problems," *Decision Analysis*, vol. 2, pp. 69–88, June 2005.

المنارة للاستشارات
www.manaraa.com

[25] Brown, B., *Introduction to Defense Acquisition Management*. Fort Belvoir, VA: Defense Acquisition University Press, tenth ed., August 2010.

[26] Browning, T. and Engel, A., "Designing systems for adaptability by means of architecture options," *Systems Engineering*, vol. 11, pp. 125–146, May 2008.

[27] Brusoni, S., Provera, S., and Prencipe, A., "Making design rules: A multidomain perspective," *Organizational Science*, vol. 17, no. 2, pp. 179–189, 2006.

[28] Calloni, B., "Mil-spec vs. cots standards: Necessary harmony for affordable multilevel secure architectures," affordable safe and secure cots software initiative, Lockheed Martin Corporation, http://sysa.omg.org/docs/, 2004.

[29] Canada, J., Sullivan, W., and White, J., *Capital Investment Analysis for Engineering and Management*. Upper Saddle River, NJ: Prentice Hall, second ed. ed., 1996.

[30] Cellini, S. and Kee, J., *Handbook of Practical Program Evaluation*, ch. 21, pp. 493–530. John Wiley and Sons, 2015.

[31] Chadda, R. and Chitgopekar, S., "A generalization of the logistics curves and long-range forecast of residence telephones (1996-1991)," *The Bell Journal of Economics and Management Science*, vol. 2, pp. 542–560, 1971.

[32] Chairman of the Joint Chiefs of Staff, *Joint Capabilities Integration And Development System - 3170.01H*, January 2012.

[33] Chairman of the Joint Chiefs of Staff, *Manual for the Operation of the Joint Capabilities Integration and Development System*, January 2012.

[34] Chen, W., Wiecek, M., and Zhang, J., "Quality utility - a compromise programming approach to robust design," *Journal of Mechanical Design*, vol. 121, pp. 179–187, 1999.

[35] Chen, W., Allen, J., Tsui, K.-L., and Mistree, F., "Management for robust design: Minimizing variations by noise factors and control factors," *Journal of Mechanical Design*, vol. 118, pp. 478–485, December 1996.

[36] Choi, J., Shin, J., and Lee, J., "Strategic management of new products: Exante simulation and market segmentation," *International Journal of Market Research*, vol. 55, no. 2, pp. 289–314, 2013.

[37] Chorpita, B., Daleiden, E., and Weisz, J., "Modularity in the design and application of therapeutic interventions," *Applied and Preventative Psychology*, vol. 11, pp. 141–156, September 2005.

[38] Chu, A., Kalaba, R., and Spingarn, K., "A comparison of two methods for determining the weights of belonging to fuzzy sets," *Journal of Optimization Theory and Applications*, vol. 27, pp. 531–528, April 1979.

المنارة للاستشارات
www.manaraa.com

[39] Cole, B., *An Evolutionary Method for Synthesizing Technology Planning and Architectural Advance.* PhD thesis, Georgia Institute of Technology, Atlanta, GA, May 2009.

[40] Conklin, J., *Wicked Problems and Social Complexty*, ch. 1, pp. 1–20. Cog Nexus, Wiley, 2005.

[41] Conrad, K., "Probability distributions and maximum entropy." University of Connecticut Course Supplement, Mansfield, CT.

[42] Consultancy, T., "Technology refresh management." White Paper, April 2011. www.tcs.com/consultancy.

[43] Cooksey, D., *A Portfolio Approach To Design In The Preseence Of Secenario-Based Uncertainty.* PhD thesis, Georgia Institute of Technology, August 2013.

[44] Cox, J., Ross, S., and Rubenstein, M., "Options pricing: A simplified approach," *Journal of Financial Economics*, vol. 7, pp. 229–263, 1979.

[45] Danner, T., *A Formulation of Multidimensional Growth Models for the Assessment and Forecast of Technology Attributes.* PhD thesis, Georgia Institute of Technology, Atlanta, GA, August 2006.

[46] Das, I. and Dennis, J., "Normal-boundary intersection: A new method for generating the pareto surface in nonlinear optimization problems," *SIAM Journal of Optimization*, vol. 8, pp. 631–657, 1998.

[47] Daskilewicz, M., *Methods For Parameterizing And Exploring Pareto Frontiers Using Barycentric Coordinates.* PhD thesis, Georgia Institute Of Technology, May 2013.

[48] Daskilewicz, M., German, B., Takahashi, T., Donovan, S., and Shanjanian, A., "Effects of disciplinary uncertainty on multi-objective optimization in aircraft conceptual design," *Structural and Multidisciplinary Optimization*, vol. 44, pp. 831–846, July 2011.

[49] de Weck, O., de Neufville, R., and Chaize, M., "Enhancing the economics of communication satellite via orbital reconfigurations and staged deployment," *American Institute of Aeronautics and Astronautics (AIAA)*, 2003.

[50] DeBecker, A. and Modis, T., "Determination of uncertainties in s-curve logistic fits," *Technological Forecasting and Social Change*, vol. 46, pp. 153–173, 1994.

[51] Defense Acquisition University, https://acc.dau.mil/, *Defense Acquisition Guidebook.*

[52] Department of Defense, *DoD Directive 5000.01 - The Defense Acquisition System*, May 2003.

المنارة للاستشارات
www.manaraa.com

[53] Department of Defense, *Risk Management Guide for DoD Acquisition*, sixth ed., August 2006.

[54] Department of Defense, *DoD Directive 5000.02: Operation of the Defense Acquisition System*, January 2015.

[55] Department of Defense Open Systems Arcitecture Data Rights Team, https://acc.dau.mil/OSAGuidebook, *Open Systems Architecture Contract Guidebook*, 1.1 ed., June 2013.

[56] Department of Defense: Systems Management College, Fort Belvoir, VA, *Systems Engineering Fundamentals*, January 2001.

[57] Dieter, G. and Schmidt, L., *Engineering Design*. New York, NY: McGraw-Hill, 4th ed., 2009.

[58] Dodson, E., "Measurement of the state of the art and technological advance," *Technological Forecasting and Social Change*, vol. 27, pp. 129–146, 1985.

[59] Domercant, C., *ARC-VM: An Architecture Real Options Complexity-Based Valuation Methodology For Military Systems-Of-Systems Acquisitions*. PhD thesis, Georgia Institute of Technology, Atlanta, GA, December 2011.

[60] Doyle, J., "B2pcoe open system architecture - final report," Tech. Rep. A2299, ACI Technologies, Inc., Philadelphia, PA, 2011.

[61] Dunbar, S., "Stochastic process and advanced mathematical finance: Prop-proper of geometric brownian motion." Electronic. www.math.unl.edu.

[62] Easingwood, C., "Product life cycle patterns for new industrial products," *R&D Management*, vol. 18, pp. 23–32, 1988.

[63] Easingwood, C., Mahajan, V., and Mueller, E., "A nonsymmetric responding logistics model for forecasting technological substitution," *Technological Forecasting and Social Change*, vol. 20, pp. 199–213, 1983.

[64] Fox, R., "Defense acquisition reform, 1960-2009," tech. rep., Center of Military History, United States Army, Washington D.., 2011.

[65] Frigo, M. and Johnson, S., "The design and implementation of fftw3," *Proceedings of IEEE*, vol. 93, no. 2, pp. 216–231, 2005.

[66] GAO, "Assessments of selected weapon programs," Tech. Rep. GAO-09-326SP, Givernment Accountability Office, 2009.

[67] GAO, "Trends in nunn-mccurdy breaches," Tech. Rep. GAO-11-295R, Government Accountability Office, 2011.

[68] GAO, "Assessments of selected weapon programs," Tech. Rep. GAO-09-326SP, Givernment Accountability Office, 2013.

المنارة للاستشارات
www.manaraa.com

[69] Gershenson, J., Prasad, G., and Zhang, Y., "Product modularity: Definitions and benefits," *Journal of Engineering Design*, vol. 14, pp. 295–313, September 2003.

[70] Gillis, M., "Open systems joint task force gets the word out," *Program Manager*, vol. 28, pp. 44–47, July 1999.

[71] Ginsburg, L., Akcakaya, R., and Ferson, S., "Where one cannot speak: When input distributions are unknown," 1996.

[72] Gonzalez-Zugasti, J., Otto, K., and Baker, J., "Assessing value in platform product family design," *Research in Engineering Design*, vol. 13, pp. 30–41, 2001.

[73] Government Accountability Office: Applied Research and Methods, *Cost Estimating and Assessment Guide: Best Practices for Developing and Managing Capital Program Costs*, March 2009.

[74] Haimes, Y., Ladson, L., and Wismer, D., "On a bicriterion formulation of the problem of integrated system identification and system optimization," *IEEE Transactions On Systems, Man, And Cybernetics*, vol. 1, pp. 296–297, January 1971.

[75] Hanratty, M., Lightsey, R., and Larson, A., "Open systems and the systems engineering process," *Acquisition Review Quarterly*, pp. 47–60, Winter 1999.

[76] Harvey, A., "Time series forecasting models based on the logistics forontier," *Journal of the Operational Research Society*, vol. 35, pp. 641–646, 1984.

[77] Hazelrigg, G., *Systems Engineering: An Approach to Information-Based Design*. Upper Saddle River, NJ: Prentice Hall, 1996.

[78] Heeler, R. and Hustad, T., "Problems in predicting new product growth for consumer durables," *Management Science*, vol. 26, pp. 1007–1020, 1980.

[79] Henning, J., "Spec cpu suite growth: An historic prespective," *Computer Architecture News*, vol. 35, no. 1, pp. 65–68, 2007.

[80] Holtta, K., Suh, E., and de Weck, O., "Tradeoff between modularity and performance for engineered systems and products," in *International Conference on Engineering Design*, (Melbourne), August 2005.

[81] Hora, S., "Aleatory and epistemic uncertainty in probability elicitation with an example from hazardous waste management," *Reliability Engineering and System Safety*, vol. 54, pp. 217–223, November-December 1996.

www.manaraa.com

[82] Horn, J., Nafpliotis, N., and Goldberg, D., "Niched pareto genetic algorithm for multi-objective optimization," in *Proceedings of the First IEEE Conference on Evolutionary Computation*, vol. 1 of *IEEE World Congress on Computational INtelligence*, (Orlando, FL, USA), pp. 82–87, 1994.

[83] Hummel, J., "National goods versus public goods: Defense, disarmament, and free riders," *The Review of Austrian Economics*, vol. 4, pp. 88–122, 1990.

[84] Humphreys & Associates, Inc., www.humphreys-assoc.com, *Basic Concepts of Earned Value Management (EVM)*, 2012.

[85] Hwang, C.-L. and Yoon, K., *Multiple Attribute Decision Making: Methods and Applications*. Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, 1981.

[86] Ibanez, A., "Valuation by simulation of contingent claims with multiple early exercise opportunities," *Mathematical Finance*, vol. 14, pp. 223–248, April 2004.

[87] Ibanez, A. and Zapatero, F., "Monte carlo valuation of american options through computation of the optimal exercise frontier," *Journal of Financial and Quantitative Analysis*, vol. 39, no. 2, pp. 253–275, 2004.

[88] INCOSE, *Systems Engineering Handbook: A Guide for System Life Cycle Process and Activities*. International Council on Systems Engineering, vol. 3 ed., 2006. INCOSE-TP-2003-002-03.

[89] Inman, O., *Technology Forecasting Using Data Envelopment Analysis*. Systems science, Portland State University, Portland, OR, 2004.

[90] Inman, O., Anderson, T., and Harmon, R., "Predicting u.s. jet fighter aircraft introduction from 1944 to 1982: A dogfight between regression and tfdea," *Technological Forecasting and Social Change*, vol. 73, pp. 1178–1187, 2006.

[91] Jackson, D., Lamartin, D., and Yahn, J., "Wami final report," tech. rep., Secretary of the Air Force, Acquisition: Information Dominance, 2012.

[92] Jain, S. and Oosterlee, C., "Pricing high dimensional bermudan options using the stochastic grid method," *International Journal of Computer Methematics*, vol. 89, no. 9, pp. 1186–1211, 2012.

[93] Jaynes, E., "Information theory and statistical mechanics," *Physical Review*, vol. 106, pp. 620–630, May 1957.

[94] Kalligeros, K., *Platforms and Real Options in Large-Scale Engineering Systems*. PhD thesis, Massachusetts Institute of Technology, June 2006.

[95] Keeney, R. and Raiffa, H., *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. Cambridge University Press, 1993.

المنارة للاستشارات
www.manaraa.com

[96] KENDALL, F., "Better buying power: Foreword," *Defense Acquisition Technology and Logistics*, pp. 1–4, September-October 2011. Available at http://www.dau.mil/pubscats/ATL

[97] KHAJAVIRAD, A., MICHALEK, J., and SIMPSON, T., "An eficient decomposed multi-objective genetic algorithm for solving the joint product family selection and design problem with generalized commonality," *Structural and Multidisciplinary Optimization*, vol. 39, pp. 187–201, 2009.

[98] KNIGHT, K., *A Study of Technology Innovation - The Evolution of Digital Computers*. PhD thesis, Carneige Institute of Technology, Pittsburg, PA, November 1963.

[99] KUJAWSKI, E., "Unintended consequences of performance specifications for the reliability of military weapon systems," *Systems Engineering*, vol. 13, no. 4, pp. 405–412, 2010.

[100] LAKE, D., "Technology, qualitative superiority, and the overstreched american military," *Strategic Studies*, vol. 71, pp. 71–99, Winter 2012.

[101] LEE, J., "Multiple-objective problems: Pareto optimal solutions by methods of proper equality constraints," *IEEE Transactions on Automatic Control*, vol. AC-21, pp. 641–650, October 1976.

[102] LEVENBACH, H. and REUTER, B., "Forecasting trending time series with relative growth rate models," *Technometrics*, vol. 18, pp. 261–272, 1976.

[103] LEVIN, H. and MCEWAN, P., *Cost-Effectiveness Analysis: Methods and Applications*. Sage Publications, Inc., 2nd ed., 2001.

[104] LICHTENSTEIN, S. and NEWMAN, J., "Empirical scaling of common verbal phrases associated with numerical probabilities," *Psychonometric Science*, vol. 9, no. 10, pp. 563–564, 1967.

[105] LIM, D., "Technology forecasting using dea in the presence of infeasibility," in *INFORMS*, 2014.

[106] LIPSON, H., POLLACK, J., and SUH, N., "Promoting modularity in evolutionary design," in *Proceedings of DETC'01: ASME Design Engineering Technical Conference*, (Pittsburg, PA), September 2001.

[107] LITTLE, A. and WOOD, K., "Functional analysis: A fundamental empirical study for reverse engineering, benchmark, and redesign," in *Proceedings of the 1997 Design Engineering Technical Conference*, no. DTM-3879 in 97-DETC, (Sacremento, CA), 1997.

[108] LONGSTAFF, F. and SCHWARTZ, E., "Valuing american options by simulation: A simple least squares approach," *The Review of Financial Studies*, vol. 14, no. 1, pp. 113–147, 2001.

المنارة للاستشارات
www.manaraa.com

[109] LOPEZ, R. and BECK, A., "Reliability-based design optimization strategies based on form: A review," *Journal of the Brazilian Society of Mechanical Science and Engineering*, vol. 34, Oct./Dec. 2012.

[110] LUCE, R. and RAIFFA, H., *Games and Decision: Introduction and Critical Survey*. Dover Publications, 1957.

[111] MAKRIDAKIS, S., WHEELWRIGHT, S., and MCGEE, V., *Forecasting: Methods and Forecasting*. New York, NY: John Wiley & Sons, second ed., 1983.

[112] MANSFIELD, E., "Technological change and the rate of imitation," *Econometrica*, vol. 26, pp. 741–766, 1968.

[113] MARK, A., YOUNOSSI, O., BRANCATO, K., BLICKSTEIN, I., and GRAMMICH, C., "Why has the cost of fixed-wing aircraft risen?: A macroscopic examination of the trends in u.s. military aircraft costs over the past several decades," tech. rep., RAND Corporation, 2008.

[114] MARKISH, J. and WILLCOX, K., "A value-based approach for commercial aircraft conceptual design," *AIAA Journal*, vol. 41, no. 10, pp. 2004–2012, 2012.

[115] MARLER, R. and ARORA, J., "Survey of multi-objective optimization methods for engineering," *Structural and Multidisciplinary Optimization*, vol. 26, pp. 369–395, 2004.

[116] MARLER, R. and ARORA, J., "The weighted sum method for multi-objective optimization: new insights," *Structural and Multidisciplinary Optimization*, vol. 41, pp. 853–862, 2010.

[117] MARSHALL, R., LEANEY, P., and BOTTERELL, P., "Enhanced product realization through modular design: An example of product/process integration," *Journal of Integrated Design and Process Technology*, vol. 3, pp. 143–150, 1998.

[118] MARTINEZ, A. M. M. and SIMPSON, T., "A penalty function for product family design using physical programming," *ASME Journal of Mechanical Design*, vol. 124, no. 2, pp. 164–172, 2002.

[119] MARTINO, J., "Measurement of technology using tradeoff surfaces," *Technological Forecasting and Social Change*, vol. 27, pp. 147–160, 1985.

[120] MARTINO, J., "A comparison of two composite measures of technology," *Technological Forecasting and Social Change*, vol. 44, pp. 147–159, 1993.

[121] MARTINO, J., *Technological Forecasting for Decision Making*. Wiley-Holland, third ed., 1993.

المنارة للاستشارات
www.manaraa.com

[122] Mavris, D. and Bandte, O., "A probabilistic approach to multivariate constrained robust design simulation," in *2nd World Aviation Congress and Exposition*, (Anaheim, CA), American Institute of Aeronautics and Astronautics, October 1997.

[123] Messac, A., "Physical programming - effective optimization for computational design," *AIA Journal*, no. 34, pp. 149–158, 1996.

[124] Messac, A., "From dubious construction of objective functions to the application of physical programming," *AIAA Journal*, vol. 38, pp. 155–163, January 2000.

[125] Messac, A. and Ismail-Yahaya, A., "Required relationship between objective function and pareto frontier orders: Practical implications," *AIAA Journal*, vol. 39, pp. 2168–2174, November 2001.

[126] Messac, A. and Mattson, C., "Normal constraint method with gaurantee of even representation of complete pareto frontier," *AIAA Journal*, vol. 42, no. 10, pp. 2101–2111, 2004.

[127] Messac, A., Sundararaj, G., Tappeta, R., and Renaud, J., "Ability of objective functions to generate points on nonconvex pareto frontiers," *AIAA Journal*, vol. 38, pp. 1084–1091, June 2000.

[128] Meyer, M. and Lehnerd, A., *The Power of Product Platforms, Building Value and Cost Leadership*. New York, NY: The Free Press, 1997.

[129] Meyer, M. and Booker, J., *Eliciting and Analyzing Expert Judgement: A Practical Guide*, vol. 5 of *Knowledge Based Systems*. Academic Press, 1991.

[130] Mistree, F., Hughes, O., and Bras, B., *Compromise Decision Support Problem and the Adaptive Linear Programming Algorithm*, ch. 11, pp. 251–290. AIAA, 1993.

[131] Murphy, R. and Birkler, J., "Dollar value and risk levels: Changing how weapon system programs are managed," tech. rep., The RAND Corporation, 2009.

[132] Muthyalampalli, R., "Implementation of fast fourier transform for image proprocess in directx 10," tech. rep., Intel, 2011.

[133] Navy, U., "Information dominance roadmap: 2013-2028," tech. rep., Department of Defense, 2013.

[134] Neuville, R., "Real options: Dealing with uncertainty in systems planning and design," *Integrated Assessment*, vol. 4, pp. 26–34, March 2003.

[135] Office of Aerospace Studies, Air Force Materiel Command (AFMC), *Analysis of Alternatives (AoA) Handbook: A Practical Guide to Analyses of Alternatives*, July 2008.

المنارة للاستشارات
www.manaraa.com

[136] Office of the Staff Judge Advocate: Space and Missile Systems Center, *Acquiring and Enforcing the Governments Rights in Technical Data and Computer Software Under Department of Defense Contracts: A Practical Handbook for Acquisition Professional*, 6 ed., March 2014.

[137] Pahl, G., Beitz, W., Feldhusen, J., and Grote, K.-H., *Engineering Design: A System Approach*. Springer-Verlag, 2007.

[138] Penfield, P., "Principle of maximum entropy." Course Notes, April 2003.

[139] Pennock, M., Rouse, W., and Kollar, D., "Transforming the acquisition enterprise: A framework for analysis and a case study of ship acquisition," *Systems Engineering*, vol. 10, no. 2, pp. 99–117, 2007.

[140] Pine, B., *Mass Customization: The New Frontier in Business Competition*. Boston, MA: Harvard Business School Press, 1993.

[141] Porter, A., Roper, A., Mason, T., and Rossini, F., *Forecasting and Management of Technology*. New York, NY: Wiley, 1991.

[142] Pratap, D., Agarwal, S., and Meyarivan, T., "A fast elitist miulti-objective genetic algorithm: Nsga-ii," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.

[143] Project, D. A. P. A., "Defense acquisition performance assessment report," tech. rep., United States Government, January 2006.

[144] Rayman, S. and Zwecher, M., "Monte carlo estimation of american call options on the maximum of several stocks," *Journal of Derivatives*, vol. 5, no. 1, 1997.

[145] Rendon, R., "Analysis of mosa implementation in navy acquisition programs," tech. rep., Naval Post Graduate School, 2010.

[146] Robinson, J., *The Economics of Imperfect Competition*. London: Macmillan, 1933.

[147] Saaty, T., "A scaling method for priorities in heirarchical structures," *Journal of Mathematcal Psychology*, vol. 15, pp. 234–281, June 1977.

[148] Saleh, J., Hastings, D., and Newman, D., "Flexibility in system design and implications for aerospace systems," *Acta Astronautica*, vol. 53, pp. 927–944, 2003.

[149] Schrage, D., "Technology for rotorcraft affordability through integrated product/process development," in *American Helicopter Society 55th Annual Forum*, American Helicopter Society, May 1999.

[150] Shannon, C., "The mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, July-October 1948.

المنارة للاستشارات
www.manaraa.com

[151] Sharif, M. and Islam, M., "The weibull distribution as a general model for forecasting technological change," *Technological Forecasting and Social Change*, vol. 18, pp. 247–256, 1980.

[152] Sharma, J., *STASE: Set Theory-Influenced Architecture Space Exploration.* PhD thesis, Georgia Institute of Technology, 2014.

[153] Simpson, T., "Product platform design and customization: Status and promise," *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, vol. 18, pp. 3–20, 2004.

[154] Simpson, T. and D'Souza, B., "Assessing variable levels of platform commonality within a product family using a multiobjective genetic algorithm," *Concurrent Engineering: Research and Applications*, vol. 12, pp. 119–129, June 2004.

[155] Smith, S., *The Scientist and Engineer's Guide to Digital Signal Processing.* California Technical Publishing, San Diego, CA, first ed., 1997.

[156] Solomon, P., "Integrating systems engineering with earned value management," *Defense AT&L*, vol. May-June, 2004.

[157] Sosa, M., Agrawal, A., Eppinger, S., and Rowles, C., "A design approach to define modularity of product components," in *ASME International Design Engineering Technical Conferences & Computers and Information in Engineering COnference*, 2005.

[158] Steffans, P. and Douglas, E., "Valuing technology iinvestment: Use of real options thinking, but forget real options valuation," *International Journal of Technoentrepreneurship*, vol. 1, no. 1, pp. 58–77, 2007.

[159] Steward, D., "The design structure system: A method for managing the design of complex systems," *IEEE Transactions in Engineering Management*, vol. 28, no. 3, pp. 71–84, 1981.

[160] Stone, R. and Wood, K., "Development of a fufunction basis for design," *Journal of Mechanical Design*, vol. 122, pp. 359–370, December 2000.

[161] Stone, R., Wood, K., and Crawform, R., "A heuristic method for identifying modules for product architectures," *Design Studies*, vol. 21, pp. 5–31, January 2000.

[162] Suh, N., *The Principles of Design.* Oxford University Press, 1990.

[163] Suh, N., "Axiomatic design theory for systems," *Research in Engineering Design*, vol. 10, pp. 189–209, 1998.

المنارة للاستشارات
www.manaraa.com

[164] Sullivan, M., "Defense acquisitions: Dod's requirements determination process has not been effective in prioritizing joint capabilities," Report to the Committee on Armed Services, U.S. Senate GAO-08-1060, Government Accountability Office (GAO), September 2008.

[165] Tanis, E. and Zimmerman, D., *Probability and Statistical Inference*. Prentice Hall, 2010.

[166] Team, N. O. A. E., "Key open sub-system (koss) tool: Koss description and application," Tech. Rep. SPR-09-674, U.S. Navy Air Systems Command (NAVAIR), August August.

[167] Thurston, D., *Decision Making in Engineering Design*, ch. 3, pp. 15–19. American Society of Mechanical Engineers (ASME), 2006.

[168] Tigert, D. and Farivar, F., "The bass new product growth model: A sensitivity analysis for a high technology product," *Journal of Marketing*, vol. 45, pp. 81–90, 1981.

[169] Triantaphyllou, E., *Multi-Criteria Decision Making: A Comparative Study*. Kluwer Academic Publishers, 2000.

[170] Tu, J., Choi, K., and Park, Y., "A new study on reliability-based design optimization," *Journal of Mechanical Design*, vol. 121, pp. 557–564, December 1999.

[171] Tversky, A. and Kahneman, D., *Judgement Under Uncertainty: Heuristics and Biases*, ch. 1, pp. 1–20. Cambridge University Press, 1982.

[172] Twiss, B., *Forecasting for Technologists and Engineers*. London, UK: Peter Peregrinus Ltd., 1992.

[173] Ullman, D., "Decisions based on analysis of alternatives," *Phalanx, Military Operations Research Society (MORS) Journal*, vol. 44, pp. 1–24, January 2009.

[174] Ulrich, K. and Tung, K., "Fundamentals of product modularity," in *ASME Issues in Design/Maunfacture Integration*, (Atlanta, GA), pp. 73–80, 1991.

[175] von Neumann, J. and Morgenstern, O., *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press, 1947.

[176] Wydler, V., "Gaining leverage over vendor lock to improve acquisition performance and cost efficiencies," Tech. Rep. 14-1262, MITRE Corporation, http://www.mitre.org/publications/technical-papers/gaining-leverage-over-vendor-lock-to-improve-acquisition-performance, April 2014.

[177] Young, P. and Ord, J., "The use of discounted least squares technological forecasting," *Technological Forecasting and Social Change*, vol. 28, pp. 263–274, 1985.

المنارة للاستشارات
www.manaraa.com

[178] YOUNG, P., "Technological growth curves. a competition of forforecast models," *Technological Forecasting and Social Change*, vol. 44, pp. 375–389, 1993.

[179] ZITZLER, E. and THIELE, L., "An evolutionary algorithm for multi-objective optimization: The strength pareto approach," Tech. Rep. 43, Swiss Federal Institute of Technology (ETH): Computer Engineering and Networks Laboratory (TIK), Zurich, Switzerland, 1998.

للاستشارات المنارة

www.manaraa.com

VITA

Eric Zellers was born June 4, 1984 in Battle Creek, MI to John Zellers and Jara Sutton. He attended Harper Creek High School in Battle Creek, MI before enrolling at the United States Military Academy in West Point, NY. After receiving a B.S. in Mechanical Engineering, Mr. Zellers received a commissioned as an Infantry officer in the U.S. Army, serving from 2006-2011. During this time he served as a Light Infantry Platoon Leader and Headquarters Executive Officer in the 101st Airborne Division, which included a 14 month deployment in support of Operation Iraqi Freedom. In February of 2011, he was assigned as a Ranger Platoon leader in the Second Ranger Battalion, where he was deployed twice in support of Operation Enduring Freedom. Mr. Zellers received an honorable discharge from the U.S. in August of 2011 and enrolled in the graduate program at the Georgia Institute of Technology, receiving a Master of Science in Aerospace Engineering in 2013. Over the past five years, Mr. Zellers has worked with a number of industry and government sponsors in the areas of satellite design, missile defense, architecture-based systems-of-systems engineering, and defense acquisitions.

المنارة للاستشارات
www.manaraa.com