A METHODOLOGY FOR THE ROBUSTNESS-BASED EVALUATION OF SYSTEMS-OF-SYSTEMS ALTERNATIVES USING REGRET ANALYSIS

A Dissertation
Presented to
The Academic Faculty

by

Benjamin Hancock Poole

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

Georgia Institute of Technology
August, 2008

A METHODOLOGY FOR THE ROBUSTNESS-BASED EVALUATION OF SYSTEMS-OF-SYSTEMS ALTERNATIVES USING REGRET ANALYSIS

Approved by:

Dr. Dimitri Mavris, Advisor
School of Aerospace Engineering
Georgia Institute of Technology

Dr. Carlee Bishop
School of Aerospace Engineering
Georgia Institute of Technology

Dr. James McMichael
Aerospace, Transportation, and Advanced Systems Laboratory
Georgia Tech Research Institute

Dr. Janel Nixon
School of Aerospace Engineering
Georgia Institute of Technology

Dr. Daniel Schrage
School of Aerospace Engineering
Georgia Institute of Technology

Date Approved: June 27, 2008

for my grandfathers,
James Houff and David Poole

ACKNOWLEDGEMENTS

The development of this dissertation has been a process that challenged me in many dimensions. Without the strong support I have received at home, it would never have been possible. Deborah Harrison deserves thanks for countless hours of patience and for her editing prowess. As the initial sounding board for many of my ideas, she helped me clarify thoughts and transition them into my work. My parents and sisters have given me a loving base from which to jump into the wide, unforgiving world of science and engineering. Without them, I would never have made the leap. And finally, Nathan, Leah, and Chris, thank you for giving me a great escape whenever I needed it.

My committee, Drs. Carlee Bishop, Jim McMichael, Janel Nixon, and Daniel Schrage have all provided valuable insight into my work. Their diverse perspectives ensured I was able to reconcile many of the views that surround the acquisition process and helped me develop positions that could be defended from many sides. They all deserve thanks for displaying great understanding as I discovered what a dissertation really meant.

Dave Brown at the Air Force Research Lab was instrumental in this work by providing funding for research work related to the Revolutionary Hunter-Killer Initial Capabilities Document. That research led to many of the ideas contained in this dissertation and gave me the opportunity to vet my ideas with those actually involved in the acquisitions process.

At ASDL, Ian Stults, Hernando Jimenez, William Engler, Kelly Griendling, Steven Tangen, and Diana Talley all provided invaluable feedback throughout the dissertation development. Santiago Balestrini in particular contributed greatly to my ability to create models in Netlogo. Drs. Pat Biltgen and Tommer Ender helped form my early thoughts on analysis in the defense acquisition process. To everyone else at ASDL, thank you for making it a great place to spend the last four years. I will miss it.

Finally, my advisor, Dr. Dimitri Mavris has served as my mentor and friend for the past five years at Georgia Tech. His challenges have never been easy, but by working through them I have discovered more about science and myself that I ever thought I would. I truly appreciate the flexibility and patience you have shown as I explored so many different ideas.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

ACRONYMS

3DOF	3 Degree-of-Freedom
AAA	Anti-Aircraft Artillery
ACAT	Acquisition Category
AGL	Above Ground Level
ANN	Artificial Neural Network
AoA	Analysis of Alternatives
ASD(C3I)	Assistant Secretary of Defense (Command, Control, Communications, Intelligence)
ASDL	Aerospace Systems Design Laboratory
BBC	British Broadcasting Corporation
CONOPS	Concept of Operations
CPI	Continuous Process Improvement
DAU	Defense Acquisition University
DoD	Department of Defense
DoE	Design of Experiments
EW	Electronic Warfare
FAA	Functional Area Assessment
FFRDC	Federally Funded Research and Development Center
FMCS	Filtered Monte Carlo Simulation
FNA	Functional Needs Assessment
FSA	Functional Solutions Assessment
HSI	Human Systems Integration
IA	Information Assurance
ICD	Initial Capabilities Document
IDA	Institute for Defense Analysis

INCOSE	International Council on Systems Engineering
IPPD	Integrated Product Process Development
IRMA	Interactive, Reconfigurable, Matrix of Alternatives
ISR	Intelligence, Surveillance, Reconnaissance
IT	Information Technology
JCIDS	Joint Capabilities Integrated Development System
LOAC	Law of Armed Conflict
M&S	Modeling and Simulation
MAIS	Major Automated Information System
MDA	Milestone Decision Authority
MDAP	Major Defense Acquisition Program
MFE	Model Fit Error
MOE	Measure of Effectiveness
MOP	Measure of Performance
MRE	Model Representation Error
MLRS	Multiple Launch Rocket System
MS	Milestone
NATO	North Atlantic Treaty Organisation
NCDC	National Climatic Data Center
NSS	National Security System
OAS	Office of Aerospace Studies
OEC	Overall Evaluation Criterion
PA&E	Program Analysis and Execution
PAF	Project Air Force
PARM	Persistent Anti-Radiation Missile
PM	Program Manager
PSA	Personnel Support Activity

R&D	Research and Development
RCD	Robust Concept Design
RCS	Radar Cross Section
RDS	Robust Design Simulation
RDT&E	Research, Development, Test, and Evaluation
RGS	Requirements Generation System
RMA	Revolution in Military Affairs
ROD	Robust Optimal Design
RPV	Remotely Piloted Vehicle
RSM	Response Surface Methodology
RSq	Coefficient of Determination
SAM	Surface-to-Air Missile
SDD	System Development and Demonstration
SEAD	Suppression of Enemy Air Defense
SNR	Signal-to-Noise Ratio
SoS	System-of-Systems
SUV	Sport Utility Vehicle
T&E	Test and Evaluation
TDS	Technology Development Summary
TLAM	Tomahawk Land Attack Missile
TRL	Technology Readiness Level
TSFC	Thrust Specific Fuel Consumption
UAV	Unmanned Aerial Vehicle
UCAV	Unmanned Combat Aerial Vehicle
UGV	Unmanned Ground Vehicle
USA	United States Army
USAF	United States Air Force

USD(AT&L)	Undersecretary of Defense (Acquisition, Technology, and Logistics)
USMC	United States Marine Corps
USN	United States Navy
UTE	Unified Tradeoff Environment

SUMMARY

The early phases of the defense acquisition process require decisions that impact the capability of the US armed forces and the allocation of billions of dollars of taxpayer resources. Because acquisition programs often stretch for more than a decade, the operational, technological, financial, and political landscapes may be very different at system delivery than when the early acquisition decisions were made. This deep uncertainty poses a wide variety of challenges to military planners and systems designer. The current revolution in military affairs and the rise of asymmetric warfare has magnified these problems by increasing the uncertainty around adversary capability and the complexity of the systems-of-systems being designed.

The current defense acquisition process suffers from a lack of the ability to compare alternatives based on their robustness in a rigorous, quantitative fashion. A survey of the literature for robustness evaluation in engineering and the current state-of-the-art in defense acquisition identified the opportunity to develop a new method applicable to the defense acquisition process.

Other disciplines were searched for methods that could be used or used with modification for robustness evaluation in the defense acquisition process. Long-term policy analysis was identified as a promising field for methods based on its similar uncertainty, magnitude of risk, and time-horizon. A cross-fertilization of two techniques from long-term policy analysis, massive scenario generation and regret analysis, was identified as having promise for addressing robustness early in the defense acquisition process. Regret, in this context, is a measure of the degree to which a system falls short

of the optimum for a particular scenario. In order to overcome the challenges associated with using these two techniques for early defense acquisition, a new methodology was developed that coupled regret analysis and massive scenario generation with surrogate modeling techniques in a parametric environment.

The hypotheses presented in this work were tested using a modeling and simulation environment based on a strike mission in Operation Desert Storm. The first experiment tested the feasibility of a Parametric Scenario Generator, used to rapidly develop and execute a large set of scenarios for concept evaluation in the FLAMES agent-based modeling and simulation environment. The second experiment developed and tested a formal mathematical definition of Global Regret, which can be used to compare concepts in the early defense acquisition process. Additionally in the second experiment, the feasibility of approximating concept regret across the plausible scenario space with surrogate models was shown. The work for the second experiment was conducted in the JMP statistical package. The third experiment, also conducted in JMP, showed the used of a Filtered Monte Carlo Decision-Making technique for navigating the regret of concepts across the plausible scenario space. Each of the experiments performed in the hypothesis testing phase supported the hypotheses of the dissertation.

The methodology is demonstrated by using an example based on the US Air Force's persistent, precision strike mission. Eight major tasks, identified from the DoD documentation of the defense acquisition process, were completed to demonstrate the application of the Global Regret Analysis Methodology in a relevant defense acquisition process. A DoE of five concepts, including concepts with distribution of tasks among platforms and swarming approaches, were evaluated using a Parametric Scenario

Generator for a time-critical-target mission. The results of this DoE were fitted with an Artificial Neural Network type surrogate model, which was then used to explore concept regret over the entire scenario space. Based on the regret landscape and the Global Regret values, two concepts were identified as promising for future investigation.

Global Regret Analysis was qualitatively compared to five state-of-the-art robustness methods using a wide variety of criteria. The Technique for Ordered Preference by Similarity to an Ideal Solution, a multi-attribute decision making technique, was used to compare the methods. Depending on the importance of the criteria, any of the methods could be found to be the "best", however, for the majority of importance weightings, Global Regret Analysis was the strongest choice.

Global Regret Analysis shows promise for application in a number of areas of defense acquisition. In the early, Pre-Milestone A Function Solutions Analysis and Analysis of Alternatives, Global Regret Analysis provides the opportunity to understand the robustness of alternatives across a wide variety of plausible futures. Because these analyses are where the majority of the design is "locked in" and occur many years before the design's fielding, it is imperative to understand how the effectiveness might change with differing operational conditions. Additionally, once a system is selected or fielded, the landscape of effectiveness relative to operational scenarios can be quickly understood because of the parametric nature of Global Regret Analysis. Because of the mathematical formulation of Global Regret, the visualization of regions of the scenario space where a concept is "best" is intuitive. This visualization also allows for the rapid identification of areas of poor effectiveness relative to the other alternatives. As information about actual operating conditions becomes available, corrective action can be taken based on this understanding of the effectiveness landscape.

CHAPTER 1

INTRODUCTION

1.1 Shift in Warfare

A major doctrine shift is underway in the United States Department of Defense (DoD). Historically, warfare has been conducted by massing forces to do battle with an enemy of comparable strength and intelligence [73]. These peer-on-peer battles were characterized by an enemy who was clearly identifiable and fought by a set of rules established by international treaty or convention. In these types of peer-on-peer conflicts, a set of fairly simple relationships can be used to determine which side will prevail [71].

Since the dissolution of the Soviet Union, however, the landscape of warfare has drastically changed. Technologies perfected in the past century have greatly magnified the potential for non-state actors to wield influence on an international scale. Consequently, the US Military must adapt its mission to fight not only the large battles of the peer-to-peer conflict era, but also battles against small groups, many of whom blend easily into the native population. These conflicts can be labeled “small wars,” and their growing importance presents a major challenge for defense planners.

The United States has engaged in small wars for nearly all of its existence. From operations in Tripoli during the First Barbary War [252] to the Boxer Rebellion in China [185] and to the operations in Bosnia in the late 1990s [32], the US has consistently deployed forces to “hot-spots” around the world. These small wars have typically garnered less attention than the major conflicts because of their expeditionary nature and small drain on national resources. In many cases these operations have been entirely prosecuted by the United States Marine Corps, which under the expeditionary model was almost completely self sufficient for the duration of many small conflicts. The expeditionary nature of these conflicts required Marines to integrate their fighting forces, land, sea, and air, and logistics support into a cohesive force that could win battles far from reinforcements or supply lines.

The integrated approach used by the Marine Corps has historically enabled the expeditionary warfare model. [232] As non-state actors become more important on the world stage, a concept closely coupled with fourth-generation warfare [245], the DoD has laid out a strategy that brings its functions closer to the Marine Corps model, currently the MAGTF. This new strategy is commonly known as the Revolution in Military Affairs (RMA). Central to the new strategy are “joint operations,” where multiple services coordinate to achieve a goal [128]. The enabling technology approach for allowing soldiers, sailors, airmen, and marines, who exist in different command structures, to effectively prosecute joint missions is the network centric battlespace.

The network-centric warfare environment offers the potential to greatly increase the effectiveness of the military System-of-Systems (SoS). By employing mixed units, the military gains the advantages of having each type of unit, but must effectively coordinate the units. These coordinated units can potentially engage enemies from small, military operations other than war, for example stability and security operations in Iraq and Afghanistan, to large scale conflicts with emerging adversaries such as China and Russia.

The benefits of network centric joint operations are not without cost, however, and the change in doctrine has forced a change in the way the DoD acquires systems. While military systems previously served only the branch by which they were purchased, systems must now operate successfully in the integrated battlespace. These systems are designed based on the capability they provide for joint operations, instead of a set of requirements. Additionally, the technology required to integrate systems into the network-centric environment has greatly increased the cost of military systems. This increased cost means that fewer weapons can be purchased, so their effectiveness must be assured. According to Soban, the system effectiveness must be a product of the mission effectiveness over a wide variety of possible scenarios, and the cost that is required to achieve that effectiveness. [213]

1.2 MDAP Failures

The RMA has led to many changes to doctrine level policies in the DoD. The acquisition process, which develops military systems while working closely with industrial suppliers, traditionally has taken many years to field new systems. Especially in the case of Major Defense Acquisition Programs (MDAPs), which may stretch for two decades or more, the RMA has greatly impacted whether programs face cut-backs or cancellations.

Two MDAPs in the past 5 years have suffered high profile cancellations. In both cases, the programs were cancelled during prototype testing, and government investment in the programs were both several billions. The Comanche helicopter and Crusader artillery system programs are discussed briefly in the following sections, and will serve as case studies for identifying new challenges in the defense acquisition process brought about by the RMA.

1.2.1 Comanche

"The RAH-66 Comanche… is the centerpiece of the Army's modernization efforts for the next decade [1995-2005]." [149]

In 1995, the Army's reconnaissance fleet was composed of over 80% Vietnam-era OH-58 Kiowa helicopters. The Army's attack fleet consisted of a significant number of AH-1 Cobra helicopters, another Vietnam-era aircraft. The age of these aircraft place significant limitations on the Army's aviation operations, especially with respect to night operations, and high-altitude, high-temperature operations. Additionally, the maintenance cost of operating several different helicopters for a mission is significant. [149]

The role of the Comanche (Figure 1) was to be eyes and ears for advancing Army units in a fast paced battlefield. Because of its emphasis on low Radar-Cross-Section (RCS), the Comanche would be able to penetrate enemy lines and coordinate attacks from other units. It addition, the Comanche would be able to engage a substantial array of targets with its air-to-air missile, air-to-ground missiles and rockets, and 20mm turreted cannon. [105] According to a Congressional Budget Office report, the Comanche "could make the total combat fleet over 30 percent more capable in 2025 than [was in 1995]." [149]

Figure 1: Sikorsky RAH-66 Comanche

The improvements in Army helicopter capability provided by Comanche did carry a high price tag. In addition to the estimated 30 billion dollars (FY1996) to acquire the size Comanche fleet scheduled in 1995, the program would come at the cost of modernizing the Army's utility helicopter fleet. The utility helicopter fleet suffered many of the same age-related issues as the reconnaissance aircraft, and was badly in need of upgrades or replacement. The Congress presented four alternatives to the existing Comanche program, in 1995, that would allow an interim update of the Army's helicopter fleet, both utility and attack. Three of the four alternatives involved the complete cancellation of the Comanche program, and the fourth involved another scale back. [149]

The Comanche program managed to survive the attacks in the mid-1990s and the first prototype flew in 2003. However, in initial test flights some serious technical issues were identified that still needed to be overcome, including “software integration and testing of mission equipment, weight reduction, radar signatures, antenna performance, gun system performance, and aided target detection algorithm performance.” [105] Despite the deficiencies, Sikorsky Aircraft and Boeing both invested heavily in infrastructure for Comanche production including a 20,000 square foot assembly facility in Philadelphia. [2]

In 2004, the DoD officially cancelled the Comanche program. Defense Secretary Donald Rumsfield’s vision for the new US forces brought large scale spending programs that were perceived to be relics of the cold war under renewed attack, and it was determined that the Comanche did not meet the needs of the future US Army. There was additional concern that Unmanned Aerial Vehicle (UAV) systems would be available before Comanche’s service date that would be able to perform the same mission at a lower cost and risk to pilots lives. [105] The budget allocated for the Comanche was shifted to upgrading existing helicopter systems, and the purchase of 800 new Blackhawk helicopters. [3] Up to the cancellation date, nearly 8 billion dollars had been spent on Comanche development.

1.2.2 Crusader

In 1991, Operation Desert Storm provided the first opportunity to test the effectiveness of many US systems in what was anticipated to be a peer-to-near peer war. In the ground phase of the campaign, the US relied on a strategy that involved very rapid movement of armor. The M1 Abrams Main Battle tank, powered by a 1500 horsepower Lycoming Textron gas turbine [85], moves at somewhere near 50kph in the dark. [56] Other highly mobile units, such as the Multiple Launch Rocket System (MRLS), were able to keep up with the M1 and offer support in engaging the Iraqi armored forces. However, a major hindrance to the operation was the performance of the Paladin Artillery system. [106]

Current US doctrine for the execution of armor battles relies heavily on indirect fire artillery for destruction of enemy units and the creation of walls-of-steel to shield friendly units that are outmatched. The use of offensive, indirect artillery fire greatly reduced the risk to forward reconnaissance and direct-fire armor units, e.g., the M1 Abrams.

Unfortunately in this scenario, the diesel powered M109 Paladin, with a maximum speed of 56kph on highway, [106] was unable to keep up with the advancing US armor line. The Paladin batteries were forced to leapfrog forward, reducing by 50% the number that could fire at any given point, and also slowing the advance of the other armor systems. Based on the shortcomings of the Paladin system on the modern, fast-paced ground battlefield, the DoD began investigating a replacement system for the Paladin in 1992. [106]

The Crusader Artillery (Figure 2) system was meant to address an entire wish-list from the Army artillery community. It featured a completely automatic loading and firing mechanism, could deliver 8 rounds simultaneously on a target, separated the crew from the ammunition storage and breach compartment, included a heavily armored crew compartment with state-of-the-art navigation and communication equipment, and had an estimated top speed of 48 kph cross-country. Initially conceived as a relatively lightweight 155mm system, once all the various components had been added to the chassis, the prototype weighed in at nearly 70 tons. This weight was prohibitive due to the necessity to air deploy the vehicle. [106] To put this weight in perspective, the Abrams tank is a similar weight.

Figure 2: Crusader XM2001 155mm Self-Propelled Howitzer [234]

As a solution to the weight issue, United Defense redesigned the Crusader as a 2 vehicle system, with a 3 man crew in each vehicle. The second vehicle was for re-supply, and featured the same crew protections as the main vehicle. The vehicles could conduct the complete re-supply mission without the crews leaving the compartment.

The majority of the Army artillery community believed that the Crusader was on track to be an effective program that would meet the needs of the Army at its planned 2008 roll out. However, when the DoD came under the guidance of Defense Secretary Donald Rumsfield, a major doctrine shift occurred which put the Crusader in jeopardy. Secretary

Rumsfield envisioned the transformation of the US military into an expeditionary force, focused more on maneuver warfare and rapid deployment. As a part of this transformation, he sought to eliminate programs that were viewed as a legacy of the cold war, namely the F-22 Raptor, the RAH-66 Comanche and the Crusader Artillery System.

In May of 2002, Secretary Rumsfield requested and analysis of alternatives for US Army artillery if the Crusader were to be cancelled. He allocated 30 days for the study, but terminated the Crusader program on 8 May 2002, before the delivery of the report. [146] He cited the crusader's incompatibility with his vision for the new US military. "This decision is not about any one weapon system, but really about a strategy of warfare in his reasoning for the cancellation of the program." [146] The crusader had cost 11 billion dollars to that point.

1.3 Judging Military Systems

According to Pinker, Smith, and Booher, it is difficult, except in hindsight, to judge to goodness of a military system. [183] Even then, however, most military systems will have only been used across a limited slice of the space of possible conditions for which they could have been used. How the system would have performed in an environment outside of that for which it was tested will never really be known.

It is difficult to find a concise list of metrics that can be used to define a good military system. However, the Navy SEALS and US Marines do publish a list of the human characteristics that make a good leader or member. These characteristics are summarized in the table below. [67], [112], [231]

Table 1: Characteristics of Marines and Navy SEALS

Judgment	Justice	Dependability
Integrity	Decisiveness	Tact
Initiative	Enthusiasm	Bearing
Unselfishness	Moral Courage	Physical Courage
Knowledge	Loyalty	Endurance
Drive	Discipline	Responsibility
Accountability	Ambition	Honor
Integrity	Flexibility	Creativity
Discipline	Learning	Winning

In order to come up with a list of high-level characteristics that might describe "good" military systems, the list of human characteristics will be examined to see if any are applicable to non-human systems. Of the elements on the list, endurance, integrity, and flexibility are the few characteristics that can relate fairly directly to non-human systems. In the human sense of the word, endurance is the ability to maintain performance over time under adverse conditions. Integrity is the quality that describes an ability to perform as expected by the standards of the military service. Flexibility is the ability to change or adapt to overcome obstacles to performance.

While these three characteristics do describe things that are considered for the evaluation of military systems, they are not usually put in these terms. The integrity of a system would usually be described as the manufactured quality of the system. The endurance of a system might be described in terms of its expected service life. And finally, the flexibility of a system could be either flexibility or robustness. Some additional characteristics that are desirable in military systems are listed below.

- High Performance
- Long Service Life
- Long Time Between Failures
- Low Training Required (Simple)
- Low Acquisition Cost
- Low Maintenance Cost
- Low Disposal Cost
- Robust

Unfortunately, it is difficult to find many more traits of "good" military systems, because, depending on their particular role in the military, the direction of desirability could change. For instance, an infantryman might claim that lightweight would be a good characteristic for a military system, but a mortarman might disagree because a heavier mortar is the more stable. Stability might be a good thing for a mortarman, but a bad thing for a high-performance fighter pilot.

1.3.1 Focus of the Dissertation

Because of the recent shift in the threats that the US Military must deal with, and the possible emergent of more peer and near-peer adversaries, this dissertation effort with focus on understanding the robustness of military systems. Defense Secretary Donald Rumsfield said in 2004 that “you have to go to war with the Army you have, not the Army you want.” [167] While the Secretary’s comment was met with great criticism [192], in many situations his statement is entirely correct. Given that defense acquisition programs often stretch for at least a decade, and that the buildup for a conflict often only lasts a few months, it is impossible to custom tailor the military’s systems for each possible engagements. However, by ensuring systems are robust to a wide variety of possible operating conditions, the need to re-tool the military for every conflict could be reduced.

1.4 Establishing a Baseline

Before brainstorming ideas for improvements to a process for the evaluation of systems-of-systems, it is important to understand the state-of-the-art in robust decision making. The author could locate little information about methods for robust decision making written prior to the 1980s. However, during the 1970s and 1980s there was a revolution in manufacturing processes in Japan and the United States, respectively. The Quality Revolution [259] brought into light the need to design systems that would maintain performance levels while being insensitive to variations in the manufacturing process. While military systems need to be robust to more than just variations in manufacturing processes, the concepts of the quality revolution can be brought to bear on more general concepts of robustness.

In general, robustness deals with the insensitivity of an aspect of a design, whether cost, performance, reliability, etc, to variations beyond the control of the designer. In structural design of a wing, for example, a "robust structural design is one that in insensitive to inaccuracies in maneuver loads… due to the use of linear aerodynamic theory." [262] Additionally, a secondary piece of information that is useful for understanding the robustness of a process or product is understanding where that process breaks down. [126]

Toward the end of the 1980s a series of methodologies emerged for addressing robustness in design. These methodologies can be broadly grouped into two areas: optimizer based methods and non-optimizer based methods. Advances in computing power and the ability to perform simulations earlier in the design process greatly aided the creation of these methods.

1.4.1 Non-Optimizer-Based Robustness Methods

1.4.1.1 Taguchi's Parameter Design - 1993

Genichi Taguchi developed parameter design as a method of mitigating the effects of noises on the variability of manufactured products. Taguchi defines a noise as a variation in "primarily… three sources: environmental effects, deteriorative effects, and manufacturing imperfections." [219] According to Taguchi, parameter design can reduce the impact of all three primary noise sources. Parameter design takes place in both product design and production process design. Taguchi also enumerates three important factors for robustness: technology readiness, flexibility, and reproducibility.

Two primary metrics are used by Taguchi for evaluating the robustness of a process. The process capability index, Cp, is defined as shown in Equation 1. The tolerance is the amount of variation in the product that is allowable to remain within specification limits. The standard deviation quantifies the variation in the product output. Taguchi does not account for any shift in the mean of the product output in the process capability index.

Equation 1: Process Capability Index

$$C_p = \frac{Tolerance}{6*\sigma}$$

The second metric used by Taguchi for evaluating the robustness of a process is the signal-to-noise ratio. The noise in this formulation is a catch-all factor that includes variations from all three primary sources mentioned above. This signal-to-noise ratio is analogous to the signal-to-noise ratio (SNR) of electronic devices, where the signal of interest should be several orders of magnitude higher than the noise to allow the signal to be distinguished.

Taguchi’s formulation for robustness has some limitations because of the linear nature of the assumptions, and the lumping of all noise factors into a single term. Additionally, fairly concrete knowledge of the system being analyzed is necessary to complete the signal-to-noise calculations. Taguchi states that “new products cannot be developed smoothly and efficiently if the technologies needed for new product development and production are not available.” [219] This requirement places a significant limitation on the use of Taguchi’s methods in the pre-technology development phase of the defense acquisition process. Additionally, according to Park, et al., because of the orthogonal array used in Taguchi’s formulation, the examination of a broad design space, which is often seen in conceptual and pre-conceptual design of military systems, is difficult. [178]

1.4.1.2 Robust Concept Design - 1995

Ford and Barkan’s Robust Concept Design (RCD) attempts to address a shortfall of Taguchi’s Parameter Design by incorporating robustness considerations earlier in the design process. The creators of RCD elaborate on Taguchi’s concept of robustness by noting that “consistency of performance of all products and production processes is importantly affected by variations in their manufacture, variations in the conditions of their use and variations in the environment in which they operate.” [90] This expansion provides an important addition to Taguchi’s concept because it brings robustness out of the manufacturing process environment discussed by Taguchi and into the environment of the product’s entire life cycle. Considering the entire life cycle is especially important for military systems because of the long life-span and the high cost to operate the systems. Brigadier General Guy Townsend underlined the long life-span of military systems when he pointed out that the US Air Force has “three generations of pilots who have flown [the B-52] -- grandfather, father, and son -- in the same family. If it lasts until 2040, five generations will have flown the same plane.” [88]

Ford and Barkan rely on the same mathematical definition for robustness as Taguchi defines for Process Capability Index (Equation 1). However, citing Fabrycky, as well as general acceptance by the engineering community, the developers of RCD identify the problem definition and concept design phases of the product life cycle as the area with greatest impact on the quality of the product (as well as that with the greatest design flexibility). The RDS "window of opportunity" for increasing the robustness of a product is shown in Figure 3.

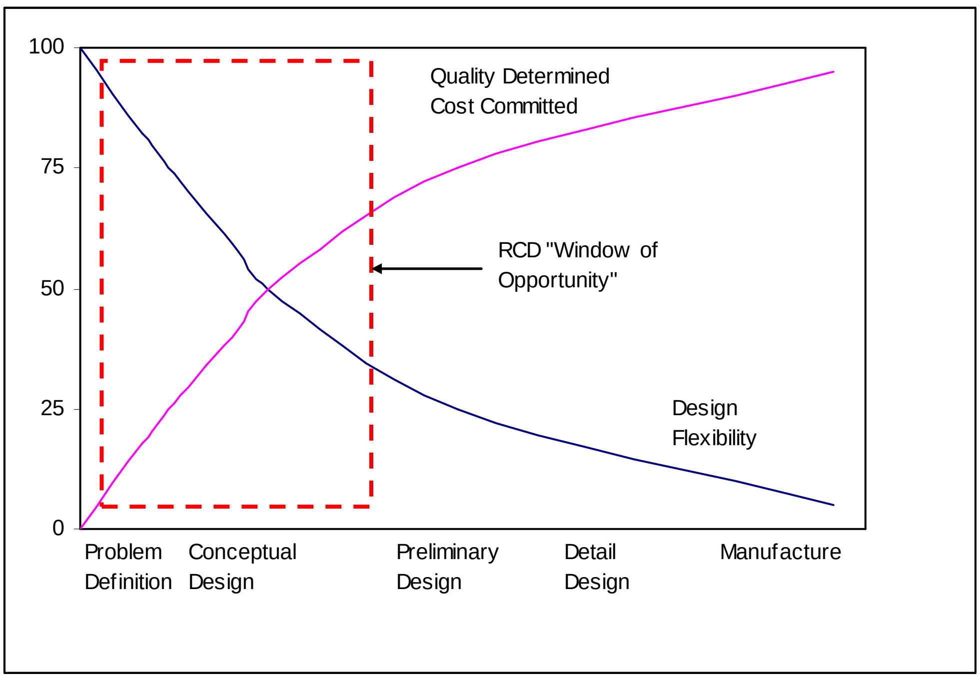


Figure 3: RCD Window of Opportunity [90]

The method for developing a robust system from the conceptual design phase using RCD consists of four stages: Definition of the Robust Problem, Derivation of Guiding Principals, New Concept Synthesis, and Concept Evaluation and Selection and Iterative Refinement. [90] RCD brings many important attributes, such as defining robustness as a primary goal, singling out and circumventing limiting constraints, however, the methodology does not address competing metrics, or significantly "outside-the-box" design. Many steps in the process refer to designer's experience, existing prototypes or products, and empirical correlations, which limit the method's applicability when prototyping or obtaining empirical data is extremely expensive (e.g. major defense acquisition) and when designers may have few experiences with the type of system being designed.

1.4.1.3 Robust Flexible Design - 1999

Kazmer and Roser developed a method for robust design with the goal of capturing both the negative effects of manufacturing variability and the potential positive effect of the manufacturing response. [131] Robust flexible design addresses two deficiencies in previous robust design methods: sensitivity to assumptions about variance in design parameters and lack of consideration of manufacturing response to flexibility. The method addresses these two areas by examining "core sources of process variation… [and] incorporates an estimate of the manufacturing response to flexibly improve the product properties during production when faces with instances of significant variation or quality loss." [131]

Kazmer and Roser suggest the following equations for evaluating the robustness for a design with multiple criteria.

Equation 2: Kazmer and Roser's Robustness

$$P_{total} = \prod_{i=1}^{n} P_i$$

$$R_i = \frac{-1}{3}\Phi^{-1} \cdot P_i \forall i$$

$$R = \frac{-1}{3}\Phi^{-1}\left(\frac{1}{2} - \frac{1}{2}\prod_{i-1}^{n}(1-2\Phi(-3R_i))\right)$$

In Equation 2, R_i is the robustness of the ith performance parameter, phi is the normal cumulative density function, and n is the number of performance parameters. This formulation allows robustness to be evaluated against several performance metrics, which may include both design and manufacturing parameters.

The methodology for implementing the above formulation requires knowledge of several factors which likely will not be available for designers working in the Pre-Milestone A phases of the Defense Acquisition Process. The method requires knowledge of product specification limits, design and manufacturing variables to a sufficient level to estimate variations in process outputs, estimates of variations in both design and process outputs and manufacturing properties. This becomes especially challenging because manufacturing design is typically not taken into consideration until much later in the Defense Acquisition Process.

1.4.1.4 Methods at the Georgia Institute of Technology - 1996 to present

Robust Design Simulation (RDS) was developed at the ASDL at Georgia Tech to address uncertainty at numerous levels of the design hierarchy. The goal in RDS is to quantify the uncertainty associated with a system and mitigate its effects. RDS defines uncertainty as the error between a mathematical model and reality, with respect to the system model, its inputs, or the operating environment. In the development of RDS effort was made to remove reliance on obsolete historical databases.

The effect of uncertainty in engineering analyses is extremely important, however, according to Mavris, "even the most elegant decomposition, approximation, and optimization schemes cannot properly account for imprecise contributing analyses, uncertain operating conditions, and ambiguous design requirements." [159] Uncertainty leads to the second criterion that is evaluated along with the objective function's expected value, the variability of the objective function.

Central to RDS is the ability to model the vehicle sizing and synthesis and the environment in which it will operate. The method first conducts ANOVA to determine the significant variables for the system level metrics. After the variable screening, response surface equations are used to model the system and Monte Carlo Analysis is conducted. The results of the Monte Carlo Analysis are viewed in the form of probability density functions (PDF) and cumulative density functions (CDF). These PDFs and CDFs are used to establish a likelihood of meeting target values. [155]

Two opportunities to improve RDS involve the inability to determine what regions of the scenario space contribute to poor system performance and providing a method for quickly comparing vastly different system concepts. Understanding the relationships between regions of the scenario space and the relative merits of different concepts is important information for decision makers. Because the capability-based analysis sought in the current defense acquisition paradigm requires consideration of many different approaches to meeting capability goals, traditional sizing and synthesis approaches are typically inadequate for the comparison. Additionally, the Monte Carlo approach of RDS suffers from some shortcomings because of the tendency of the distributions to ignore corner cases because of the Central Limit Theorem. This can become particularly problematic if the scenario space is characterized by an intelligent adversary trying to drive the scenario to technological or tactical limits.

In addition to the work at the ASDL, the System Realization Laboratory in the School of Mechanical Engineering at Georgia Tech has been addressing the design of robust systems through the use of "Families of Systems." Families of systems are developed on a common baseline model that has been designed such that the system may be continuously developed and improved to create several generations of systems that meet increasingly demanding constraints. [114] This approach is somewhat analogous to the spiral development approach favored by defense acquisitions prior to the RMA.

The majority of the work within the systems realization laboratory has focused on the product-process design and creating families that are robust to evolving constraints within that process. The techniques include the use of Design of Experiments and Response Surface Approaches, a form of surrogate models. The work has not, however, focused on SoS or the impacts of changing environmental factors after the product's manufacture.

Much of the work of the families of systems approach is focused on improving the flexibility of the systems for later adaptation to changing constraints. This approach is as opposed to choosing a system initially that's effectiveness will be insensitive to changing constraints. Both approaches have the ability to produce a "robust" solution, however, the families of systems approach requires some amount of redesign as the changes in constraints occur. These two approaches, flexibility in the design process versus selecting an insensitive design initially, represent the main two approaches to selecting a design that maintains effectiveness over a wide variety of plausible futures. However, the selection of the insensitive concept provides advantages after the system has been fielded.

While the design of flexible systems is desirable in the design phases of the system, once the system has been fielded it has potential drawbacks for military systems. One feature of the RMA is the shift toward expeditionary warfare, where supply chains are long or forces must self sustain. In those cases, flexible systems that require additional components have the potential to create logistics problems for the forward forces. The flexible components must be stored, transported, maintained, and implemented, and all of those tasks require additional manpower and expense. This approach is counter to the desire to increase the "tooth-to-tail" ratio of the armed forces. Systems that are insensitive to changes in environmental and usage conditions, however, do not require this additional logistics footprint.

1.4.2 Optimizer-Based Robustness Methods

1.4.2.1 Wilde - 1992

Wilde expands Taguchi's Parameter Design method by introducing the idea of a "quality margin" that allows the designer to use a mathematical optimizer to drive robustness. Wilde's method incorporates quality into the constraints of the optimization problem, allowing multiple quality and performance factors to be considered simultaneously. [255] Wilde uses the following definition of the quality margin.

Equation 3: Quality Margin

$$q^{\#} = \frac{Y^{\#} - y^{\#}}{Y^{\#} - T}$$

$$q_{\#} = \frac{y_{\#} - Y_{\#}}{T - Y_{\#}}$$

In Equation 1, the superscript # corresponds to the upper margin, while the subscript # to the lower. *T* is the target value for the design, Y is the average value, and y is the sample value. The quality margin is the difference between the specification limit and the largest possible deviation from target, normalized. Therefore, perfect quality, with zero standard deviation on the process and a mean on target, would be expressed as 50/50. [255]

Wilde's model relies on the assumption that a mathematical constrained optimization problem can be created for the problem at hand. For relatively simple products and processes this is likely the case, however, modeling systems-of-systems in this manner may not be possible. Additionally, multiple design objectives must be treated as constraints rather than as objectives. This creates problems when trying to address the affordability of a system, where capability and cost must be simultaneously evaluated with other, sometimes qualitative, objectives.

1.4.2.2 Robust Optimal Design - 1994

The Robust Optimal Design (ROD) was developed by Lewis and Parkinson to understand how variability in input parameters and design variables impacts a design. This type of study is also known as sensitivity analysis. Lewis and Parkinson expanded on the work of Emch and Parkinson [77] from worst-case tolerance analysis to statistical tolerance analysis. The use of statistical analysis instead of worst-case analysis allows the engineer to understand the impact of the probability of the variation in the variance of the final product. This is especially important when a large number of products will be manufactured.

ROD is based on a standard-form, multi-objective optimization problem with a number of constraints. Lewis and Parkinson determined that a linear tolerance model was inadequate for modeling problems where the skewness of product within design tolerances is significant. To address the skewness issue, the method uses a second-order tolerance model to solve the optimization problem. [145]

A limitation of the ROD method is that it relies on the use of differentiable mathematical models. Additionally, because of the assumption that parameters will be modeled with normal distributions about some mean, infeasible combinations of inputs can result, especially near the "tails" of the normal distribution. For example, if the designer was using an aircraft's cargo weight as a parameter that may have a distribution (payload weight *will* vary in military operations), it is mathematically possible to have a negative payload weight from the edges of the normal distribution, even though that value is not physically realizable.

1.4.3 Criteria for a Systems-of-Systems Robustness Evaluation Method

The existence of several methods for handling robustness during system design implies that criteria need to be established to compare the methods. The goal in the comparison is to determine if any existing method suffice for robustness analysis in the early stages of military system-of-systems design. The following table of metrics was constructed based on a decomposition of characteristics of major defense acquisition programs.

Table 2: Robustness Method Metrics

Metric	**Reasoning**
Applicability in Pre-Conceptual Design	The earliest phases of the design process are where the majority of the product quality is determined
Applicability in Conceptual Design	The earliest phases of the design process are where the majority of the product quality is determined
Robustness Evaluation at Capability Level	Capability based acquisition
Applicable to Systems-of-Systems	The revolution in military affairs has shifted design and acquisitions to networks-of-systems
Applicable to Multi-Objective Problems	Military acquisitions are inherently multi-disciplinary designs that must meet many different objectives
Applicability to Revolutionary Concepts	The emphasis on technology incorporation into military systems brings many designs outside the realm of historical or empirical design
Robustness Evaluation Based on Full Life Cycle	The long, expensive lifespan of military systems (e.g. the B-52) means that the majority of costs are not spent in the design and manufacture
Mathematical Definition of Robustness	Military acquisitions emphasize quantitative analysis wherever possible (Reference DoD 5000)
Optimizable	A robustness term that can be optimized allows Multi-disciplinary optimization to include robustness in with performance and cost
Automated	Reduction in design and re-design time is critical if many scenarios are to be studied

1.4.4 Comparison of Existing Robustness Methods

The following figure (Figure 4) displays the author's assessment of each of the five robustness evaluation methods discussed earlier with respect to the metrics in Table 2.

Poor - ●
Moderate - ⊙
Good - ○

	Not Optimizer Based			Optimization	
	Kazmer and Roser (1999)	Ford and Barkan's Robust Concept Design (1995)	Taguchi's Parameter Design	Lewis and Parkinson (1994)	Wilde (1992)
Mathematical Definition of Robustness	○	○	⊙	⊙	○
Applicability in Conceptual Design	⊙	○	⊙	●	⊙
Applicability in Pre-Conceptual Design	●	⊙	●	●	●
Robustness Evaluation at Capability Level	●	●	●	●	●
Applicable to Systems-of-Systems	⊙	⊙	⊙	⊙	●
Applicable to Multi-Objective Problems	○	⊙	⊙	⊙	●
Optimizable	○	⊙	⊙	○	○
Automated	●	●	●	○	○
Applicability to Revolutionary Concepts	●	⊙	●	●	●
Robustness Evaluation Based on Full Life Cycle	⊙	○	●	⊙	●

Figure 4: Robustness Methods

1.5 Need for a New Method

The results of comparing the robustness methods to the evaluation metrics (Figure 4) show that no existing method is entirely suited for application to early phases of military systems-of-systems evaluation. It is also apparent from the evaluation that Robustness Evaluation at the Capability Level, Applicability to Revolutionary Concepts, and Applicability in Pre-Conceptual design are areas where challenges may exist for applying a robustness methodology.

Based on the assessment of existing robustness methods, the author asserts that a new method may be able to better evaluate robustness for military systems-of-systems. This assertion will become the focus of this dissertation.

1.6 Summary

The current RMA has placed an increased burden on the defense acquisition community. The community must now acquire capabilities through the SoS they design and purchase, and those capabilities must be robust to a wide variety of possible adversaries and operational conditions. Additionally, the SoS must be able to operate effectively in the joint operations environment, placing additional constraints on the designers.

Two major failures of defense acquisition programs, the US Army's Comanche Helicopter program and the US Army's Crusader Artillery program, cost taxpayers billions of dollars with little useful military gain. The failure of these two programs underscores the need to effectively design systems that will be robust to changing battlefield conditions and adversary sets. Robustness is only one of many characteristics that can be used to judge military systems, but was selected because of the importance it plays across all of the service branches. It is also important to note that the success of a military system can only be truly judged in hindsight and that, even then, the system will have been used in only a narrow set of the possible and plausible battlefield conditions.

The current state-of-the-art methods of robust design fall into two groups: optimizer-based approaches and non-optimizer-based approaches. The methods were compared on the basis of applicability in pre-conceptual and conceptual design, robustness evaluation at the capability level, applicability to SoS, revolutionary concepts and multi-objective problems, their use of a mathematical definition for robustness, and finally the ability to optimize and automate the method. Based on the assessment of existing robustness methods, the author asserts that a new method may be able to better evaluate robustness for military systems-of-systems.

CHAPTER 2

PROBLEM DEFINITION

2.1 Introduction to Terms

Before beginning an effort to create a new methodology for making decisions based on robustness, it is important to have a common understanding of the meaning of terms that will appear throughout this dissertation. Only broad, overarching concepts will be addressed in this section, more specific concepts will be defined when they first appear in the methodology. The section will begin by defining the terms associated with the title of the dissertation: A Methodology for the Robustness-Based Evaluation of Systems-of-Systems Alternatives Using Regret Analysis.

2.1.1 Preference of Definitions

Because the focus of this work is on military acquisitions, the DoD definition of terms will be preferred, unless it is found insufficient or non-existent. In these cases, the definition used by individual branches of the US military will be used. If multiple definitions exist across branches, the most suitable will be chosen. If definitions can not be found in the DoD or across the service branches, professional association or academic definitions will be used.

2.1.2 Robustness

The concept of robustness has been discussed at length earlier in this document, including mathematical definitions that exist in the scientific and engineering literature. The DoD and service branches do not offer a natural language definition for robustness, but refer to it in many documents. Robustness has many connotations in natural language, but the particular one of interest for this dissertation states that robust means “capable of performing without failure under a wide range of conditions.” [4] The need to perform under a wide range of conditions is relevant for the vast majority of military system. Robustness will be given a mathematical definition later in this dissertation, but the natural language definition of robustness will be **the ability of a system to perform over a wide range of conditions**.

The basic idea of robustness is to be insensitive to changes in a condition around which there is uncertainty. In comparing optimal designs versus robust designs, it is expected that optimal designs will do better than robust designs in the conditions for which they were optimized, but worse than the robust designs in conditions far from those for which they were optimized. Figure 5 and Figure 6 show a general and specific example of how robust solutions can be compared to optimal solutions. In Figure 5, the effectiveness measure is plotted on the vertical axis, while the scenario variable with uncertainty is plotted on the horizontal axis. The peaked, black line might represent an optimal solution, while the more flat, red line might represent a robust solution to the same problem. The effectiveness of the red line is less than that of the black line for the peaked region, but over the rest of the range of the scenario variable the red line is more effective. The choice between the red and black solution would then depend on the likelihood of different values for the scenario variable.

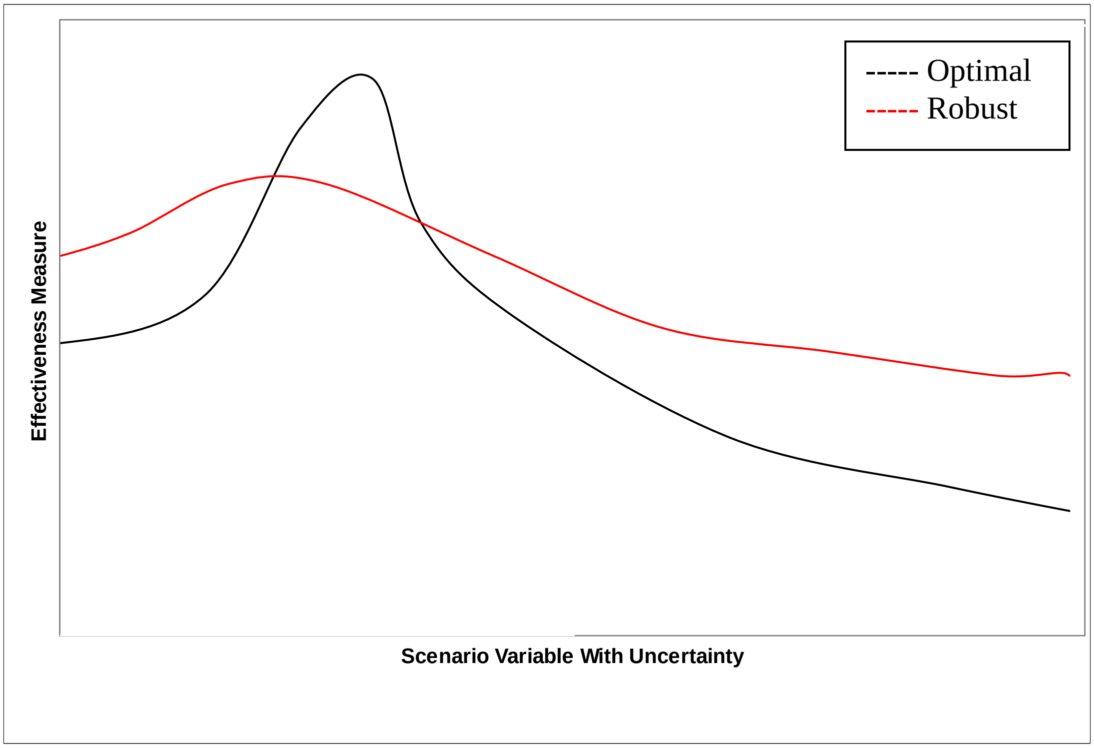


Figure 5: Allocation of Axes for Robustness

In Figure 6, a notional agricultural example has been provided. The expected yields for three crops, agave, olives, and rice, have been shown as a function of the rainfall they receive in the summer months. As can be seen in the figure, agave, a succulent, does well when rainfall is below 6 inches, but poorly above that amount because of root rot. Olives do well for a range of 5 inches to 13 inches of rain, but not above or below those amounts because of their Mediterranean evolution. Finally, rice does well in very wet climates because of its need for standing water. If the amount of rainfall was known, or subject to very little variability, a farmer would be able to select a single crop to maximize his or her yield. This would be an optimal planting for the farmer. However, if the amount of rainfall was unknown or subject to great variability, planting a mixed crop would ensure that no matter what amount of rain fell, at least some harvest would be obtained. This would be a robust planting for the farmer. Two robust mixtures are shown as the blue dashed and green lines in the figure.

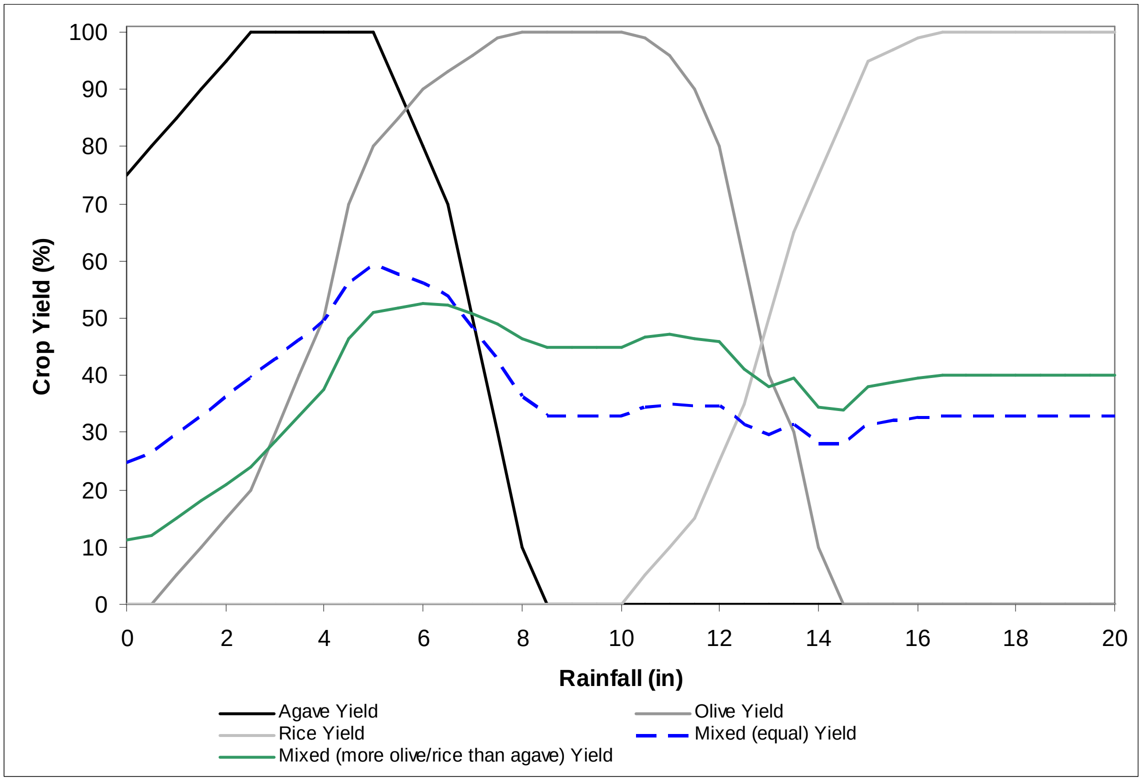


Figure 6: Notional Performance of Crops

2.1.3 Evaluation

The DoD defines evaluation, "in intelligence usage, [as the] appraisal of an item of information in terms of credibility, reliability, pertinence, and accuracy." [236] The terms, credibility, reliability, pertinence, and accuracy, however, do not fully capture the nature of evaluation that is desired for system-of-systems design problems. The individual services do not offer appropriate definitions of evaluation beyond specific applications of the term. Therefore, the definition of evaluation used in this dissertation will stem from a more general source.

Webster's definition, which is both more general and more applicable to the concept of this dissertation, is "to determine the significance, worth, or condition of usually by careful appraisal and study." [4] The determination of worth is the fundamental task of weighing different approaches and alternatives in the early phases of the DoD acquisition process. The worth is a collection of all the benefits and costs associated with a system over its life cycle. The particular definition and treatment of benefits and costs will typically be problem specific, but the general concept in the early phases of defense acquisition is to maximize a ratio of benefits to costs. The definition of evaluation for this dissertation will be a modification on Webster's definition.

Evaluation is the determination of worth (usually the ratio of benefits to costs), through careful appraisal and study.

2.1.4 System-of-Systems

In the Joint Capabilities Integrated Development System (JCIDS) documentation, the DoD defines a "set or arrangement of systems that are related or connected to provide a given capability" [203] as a system-of-systems. As Biltgen [24] observes, however, a system-of-systems is in and of itself a system. Systems-of-systems, therefore, must be a subset of systems in general, and depends on the perspective of the individual describing the system. For example, to a Federal Aviation Administration planner, the national passenger aerospace infrastructure is a complex set of systems that gives the nation the capability to move people rapidly across the continent. However, to a government transportation planner, the aerospace infrastructure is one system within the overall transportation network, which is composed of automobile transportation, rail transportation, shipping, aerospace, etc.

Because of this dependence on perspective, the author will carefully define the systems that compose the systems-of-systems discussed in this dissertation, as well as the systems-of-systems themselves. Before discussing systems-of-systems further, a definition of system will be presented.

2.1.4.1 System

The DoD defines a system as "a functionally, physically, and/or behaviorally related group of regularly interacting or interdependent elements; that group of elements forming a unified whole." [129] This definition does not, however, address a key aspect of a system: systems are created for a purpose. The International Council on Systems Engineering (INCOSE) provides the following definition in the Systems Engineering Handbook: [123] "A system is a combination of interacting elements organized to achieve one or more stated purposes." The INCOSE definition includes the fact that systems are created to do something, but does not contain the detail about possible relationships provided by the DoD. For this dissertation, the DoD definition will be expanded to include the INCOSE reference to a system's purpose.

A system is a functionally, physically, and/or behaviorally related group of regularly interacting or interdependent elements organized to achieve one or more stated purposes.

This definition is consistent with the United States Air Force (USAF) Acquisition Community definition of a system. [172] By carefully defining the meaning of a system for this dissertation, the applicability of the methods developed herein can be more effectively determined. If a particular application in inconsistent with the definitions used, addition effort will be required to determine if the method is suitable.

2.1.4.2 Defining System-of-Systems

The DoD definition of a system-of-systems gives a good starting point for determining definition that should be used in this dissertation. However, the Department of the Navy expands on the basic DoD definition by adding “the loss of any part of the system will degrade the performance or capabilities of the whole.” [64] This addition to the definition implies that not only do the systems interact when providing a capability, but they are also interdependent when providing that capability. The INCOSE definition of a system-of-systems also alludes to this interdependence by stating that the systems alone can not produce the same results. [123]

Biltgen cites five characteristics compiled by Maier [24], [150] as critical distinctions for systems-of-systems: emergent behavior, evolutionary development, operational independence of the elements, managerial independence of the elements, geographic distribution. Biltgen identifies the emergent behavior as the primary purpose of the system, which follows logically from the INCOSE definition. The two characteristics relating to the independence of the systems within the system-of-systems, managerial independence and operational independence, indicate that the systems are useful without the system-of-systems and are sometimes used independently of the system-of-systems.

Biltgen observes that the geographic distribution of the system implies that only information can be readily transferred between elements, not mass or energy. However, the author would counter by suggesting that the USAF’s refueling fleet and strike fleet are independent systems, yet the refueling fleet transfers energy (in the form of mass) to the fighters during certain system-of-systems operations.

Evolutionary development of a system-of-systems has been true for to this point. However, the newest systems-of-systems that are being developed for the US armed forces are being created from simultaneously developed systems. The Army's Future Combat System [14] is an example of a system-of-systems where nearly every system has been developed simultaneously.

2.1.5 Regret

The basic concept of regret for this dissertation is similar to the natural language usage, which relates to a sense of loss. [4] However, to be able to use regret in a rigorous way it must be quantified, and therefore a loss must be relative to something. Because regret is implicitly negative, it is desirable to eliminate the possibility of "negative regret," which would be a double negative. Therefore, the baseline to which regret will be measured is the best possible outcome in the particular scenario. The definition of regret for this dissertation incorporates that concept.

Regret is the difference between a system's evaluation metric(s) and the best system's evaluation metric(s) for a scenario.

The origins of regret analysis and the justification for its use in this dissertation will be discussed in the methodology section. Additionally, a more formal mathematical definition will be presented in the methodology section.

2.1.6 Analysis

The DoD and service branches do not offer formal definitions for analysis. According to Webster's Dictionary, analysis comes from Greek roots that mean to "break-up." In mathematics, analysis is the "systematic study of real and complex-valued continuous functions." [191] Webster also offers analysis as "an examination of a complex, its elements, and their relations." [4] The mathematical definition and natural language definitions seem to offer two different views on analysis. In mathematics, analysis is a very specific branch of mathematics (of which calculus is a part), whereas in the natural language analysis means to examine something by sectioning it in to sub-sections.

For this dissertation, analysis will follow the natural language definition, but with one addition from the mathematical definition. In studying problems for the defense acquisition process, it is important to be systematic. A systematic approach provides several advantages in this context. First, because of the transitional nature of the uniformed side of the DoD acquisitions community, a systematic approach with thorough documentation allows some continuity for the study, even as personnel change. Secondly, systematically studying acquisitions problems ensures that each problem receives consideration as rigorous as all others, or if not, the variation in rigor is justified. Finally, employing a systematic approach means that a custom, ad-hoc methodology does not have to be developed and tested for each problem; the success of the systematic method in previous applications builds credibility.

Building on the Webster natural language definition, for this dissertation, analysis is **the systematic examination of a complex by considering its elements and their relations.**

2.2 Systems Acquisition in the DoD

Use of robustness as a criterion for selecting military systems-of-systems requires understanding where in the defense acquisition process decisions relating to robustness are made. The "window of opportunity" for system design indicates that addressing robustness as early as possible in the acquisition process would be desirable. This concept is further illustrated in Figure 7, but with the addition of a knowledge curve. As Mavris shows in Figure 7, cost committed and design freedom are inversely proportional; therefore, by narrowing our system to a single concept, we have locked the majority of the design freedom and committed the majority of the cost. [158]

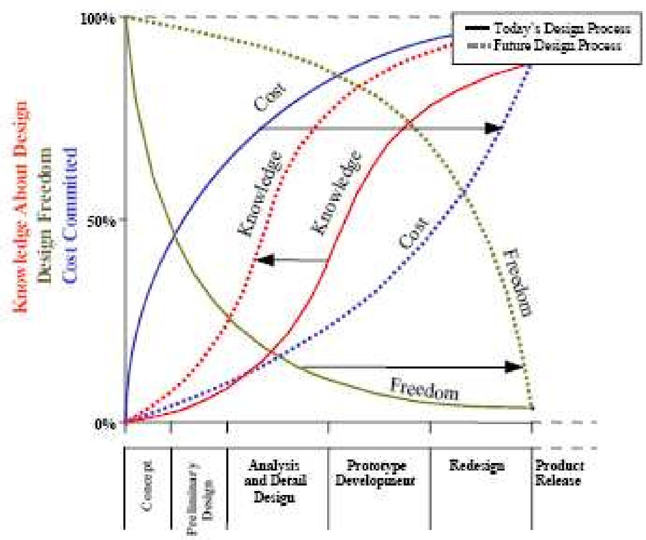


Figure 7: "Cost-Knowledge-Freedom" Shift [158]

Addressing robustness early in the acquisition process should aid in the shift of the "knowledge" curve to an earlier phase of the process, when there is more ability to change the design and less cost committed. In order to identify areas in which robustness should be considered, the following sections review the current defense acquisition process.

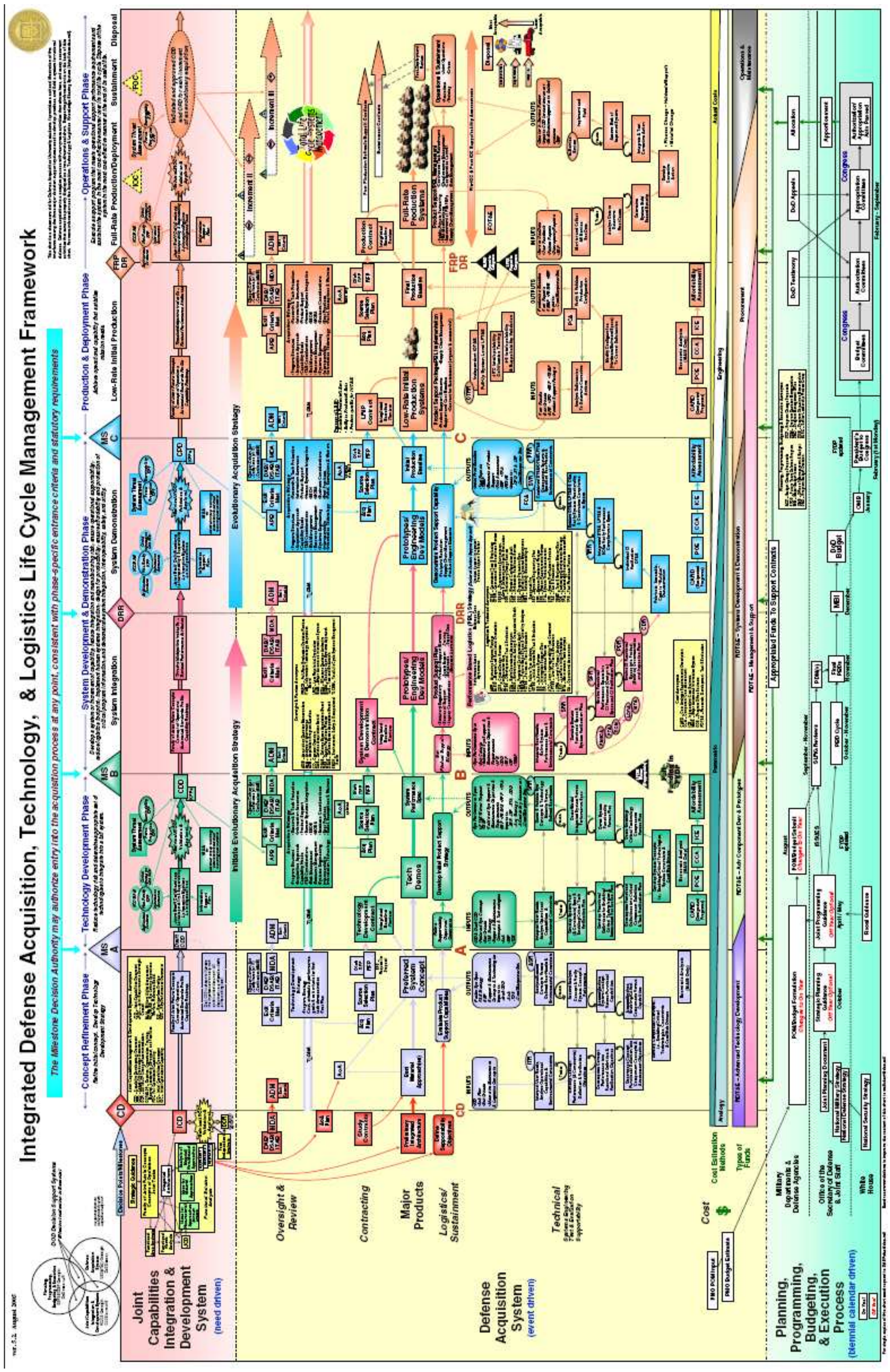


Figure 8: Complete Defense Acquisition Process [58]

2.3 Integrated Acquisition Process

Figure 8 shows the Department of Defense's Integrated Defense Acquisition, Technology and Logistics Life Cycle Management Framework. The figure shows the interaction of the Joint Capabilities Integration and Development System (JCIDS), the Defense Acquisition System, and the Planning, Programming, Budgeting, and Execution (PPBE) process. These three major areas correspond to the pink, yellow, and aqua rows in the figure, with major activities for each of the areas shown in the boxes within the rows. The events between major acquisition events are shown by color coding the boxes according to Table 3.

Table 3: Defense Acquisition Color Coding

Timeframe	**Color**
Pre-Milestone A, Pre-Concept Decision	Red
Pre- Milestone A, Concept Refinement Phase	Purple
Pre-Milestone B, Technology Development Phase	Green
Pre-DRR System Integration Phase	Pink
Pre-Milestone C, System Demonstration Phase	Blue
Post Milestone C, Production and Deployment Phase Operation and Support Phase	Orange

As shown in Figure 8, the Integrated Defense Acquisition Technology and Logistics Life Cycle Management Framework involves many analyses, which occur at many levels of the government and industry contractors. However, by looking at major components and processes first, the acquisitions process can be more easily understood.

Three interrelated tasks occur throughout the process: JCIDS, the Defense Acquisition System, and PPBE. JCIDS can be thought of as essentially *what the military or government needs*. The Defense Acquisition System is *developing the system to fulfill the needs*, and the PPBE is *how to pay for that system*. The government and industry then work together, from left to right on Figure 8, to work through major milestones, or design reviews, before the system is actually delivered to the user. There are six of these milestones or design reviews, which break the acquisition process into seven distinct phases.

2.3.1 JCIDS

JCIDS, which replaced the Requirements Generation System (RGS) in 2003, [58] defines the capabilities needed by the military or government and how systems designed to address capability gaps are to be evaluated, and is common to all branches of the US Military. [49] In a hierarchical systems decomposition [96], the capability level is the highest level objective, to which all other levels contribute. The initial steps of the JCIDS process generally precede the initiation of the Defense Acquisition System or the PPBE. The JCIDS continuously updates its information throughout the Integrated Defense Acquisition Technology and Logistics Life Cycle Management Process and exchanges information with the Defense Acquisition System and the PPBE. The JCIDS, in the early phases of the management process, works toward creating the Initial Capability Document (ICD), which guides most of the early efforts.

There are four steps in the JCIDS methodology: Functional Area Analysis (FAA), Functional Needs Analysis (FNA), Functional Solutions Analysis (FSA), and Post Independent Analysis. This analysis provides a picture of military needs for capabilities, due to gaps in current capabilities or emerging needs, and provides approaches to fill those capability needs. An emphasis is placed on considering the capability in terms of the joint operating environment. [58] A detailed view of the JCIDS process that leads to the development of the ICD is shown in Figure 9.

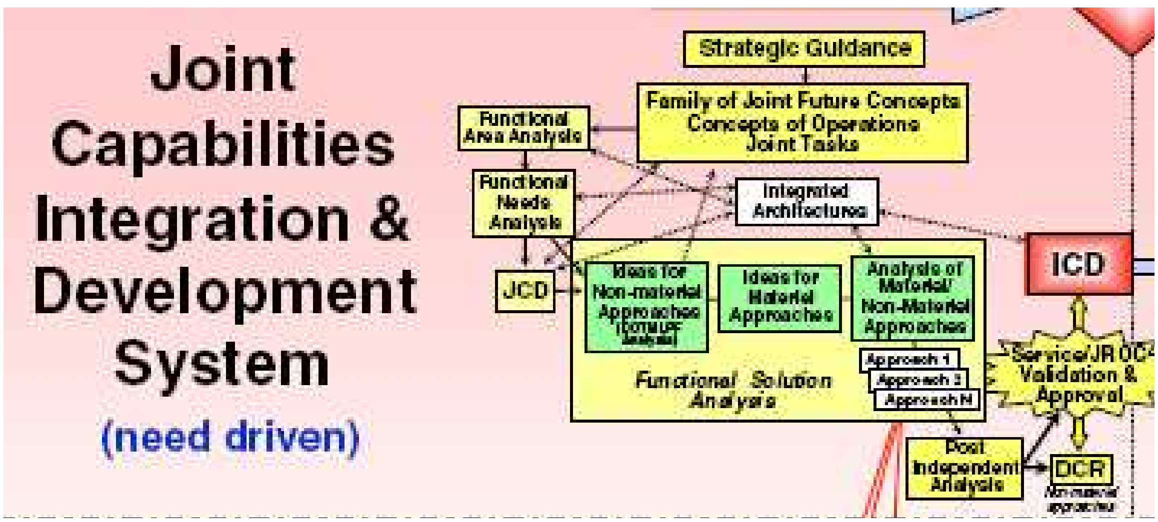


Figure 9: JCIDS Process [58]

2.3.2 PPBE

PPBE provides the funding for the development and acquisition of new military systems. Because this function involves the Do D, the White House, and the Congress, it is driven primarily by the government fiscal cycles. While capability needs identified by the JCIDS drive the Defense Acquisition System, which in turn provides an estimate of the funds needed to design and procure the system, the PPBE group, by holding the purse strings, has final control over the project. The Defense Acquisition System attempts to provide cost information to the PPBE group by first using analogy and parametric studies, the transitioning to engineering estimates, followed by the actual procurement costs. As a result, the true Life Cycle Cost of the project emerges as a spiraling development of progressively higher fidelity analyses, which are finally replaced by the cost of the fully developed and purchased system.

2.3.3 The Defense Acquisition System

This section gives an overview of the Defense Acquisition System, as outlined in the Defense Acquisition Guidebook [57], a publication of the Defense Acquisition University (DAU). The DAU is an organization within the DoD created in 1990 by the Defense Acquisition Workforce Improvement Act to better educate members of the Defense Acquisition Community. It provides guidance to all branches of the DoD with training courses in nearly all areas of the acquisition process. [59] The DoD Acquisition process is outlined below in Figure 10. In summary, the service User Needs and Technology Opportunities, as defined by the DoD in conjunction with industry technologists, feed into the initial three stages of the acquisition process. These stages begin with a refinement of the concepts identified by the DoD and technologists, which is followed by a period of technology development. Then the system goes through the actual RDT&E required to design, prototype and test the system. The system is then produced according to the specification of the System Development and Demonstration phase and transitioned to the forces acquiring the system. The final phase is the operation and support of the fielded system. The detailed workings of each phase of the process are beyond the scope of this document; however, a summary for each phase appears in the following sections.

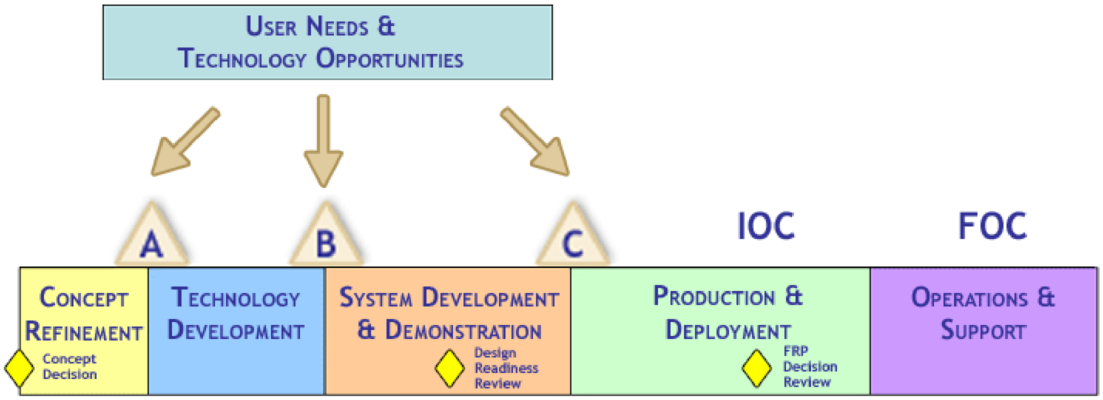


Figure 10: Defense Acquisition Process [57]

2.3.3.1 User Needs and Technology Opportunities

For more detailed information on the User Needs and Technology Opportunities section reference section 3.4 of DoD 5000.2. This phase of the Defense Acquisition Process allows for the interaction between planners at the DoD, who are aware of military needs, and technologists and industry representatives, who are aware of developing relevant technologies. It roughly corresponds with the JCIDS process, but is primarily associated with the early phases of the JCIDS. It is a parallel effort that must occur before any other phases, but is iterated upon based on the results of the milestone reviews.

In the User Needs and Technology Opportunities phase, DoD planners are tasked with defining desired capabilities for directing the process of acquiring affordable system solutions. The Initial Capabilities Document (ICD), created by the DoD, provides the foundation for the initial system development investigations. Technologist and industry representatives are tasked with identifying relevant technologies across a broad range of sources. While identifying possible technologies, they must ensure that the possibility for future competing contracts is not eliminated. [230] In short, the government is ensuring that technologists do not identify only their own technologies, therefore eliminating the chance that they would have to compete for participation in the program. Such action by the technologists could possibly reduce the performance or cost-effectiveness of the final system solution and is explicitly not allowed.

The two key phrases in the DoD's task are "defining capabilities" and "affordable" solutions. Capability based design is a relatively new concept in the DoD that relates a system design directly to its addition of an ability for the military to successfully complete some action. It emphasizes a top-down approach to design. [72] Affordable is defined in the aerospace systems design field as a ratio of a system's performance to the total life-cycle cost of acquiring, operating, maintaining and disposing of the system. [154]

2.3.3.2 Concept Refinement

For more information on the Concept Refinement phase reference section 3.5 of DoD 5000.2. Concept refinement occurs directly after the approval of the ICD, which is mandatory for the program to continue. Concept refinement is specifically the conduct of an Analysis of Alternatives (AoA), which is planned for in the ICD, and the development of the Technology Development Strategy (TDS). The AoA functions as a systematic analysis of the possible alternatives for meeting the requirements of the ICD. This AoA takes place before the initiation of any actual acquisition program, and specifically "refine[s] the selected concept documented in the ICD." [230] The AoA is expected to focus on the risks, impact and expected maturation of critical technologies, and provide information for the TDS, a major item in Milestone A.

In the AoA, the conceptual design space is reduced from a field of billions of possible system solutions [80], to a single, or in rare cases a very small group of, system alternative(s). This process occurs in a relatively short timeframe, usually less than a year, and must allow decision makers to determine that the concept selected will meet the needs of the military, be affordable, and be reasonably close to some version of a "best" solution. This process, unless conducted systematically, has the potential to leave out an acceptable level of analysis for many concepts, and can easily obscure the logic for filtering candidate system designs. Therefore, much scrutiny must be given to the methodology to ensure soundness.

The TDS, as the title suggests, provides a projected assessment of the ways in which technologies identified in the AoA, and relevant to the ICD, will mature. Specifically, the DoD is concerned with the nature of an evolutionary approach to the system maturation or the possibility of a non-spiral development. The TDS is expected to include estimates for the entire technology Research and Development (R&D) effort, including costs, timelines, and testing plans. [57]

Accurately predicting how technologies will mature is difficult, especially when they are in the early stages of development. In many cases, to effectively judge the impact of a technology, much more information is needed than is available at the current development stage. It is therefore important that the TDS accounts for the possibility of mature technologies providing a different impact than expected in the early conceptual design of the system-of-systems.

2.3.3.3 Milestone A

DoD 5000.2 outlines a set of requirements for initiation of the Technology Development phase that occurs after concept refinement (Figure 10), as well as an additional set of requirements for ship acquisitions. These requirements are broken down into two categories: statutory requirements and regulatory requirements. Some requirements are specific to Major Automated Information Systems (MAIS) acquisition, and will not be discussed in this document.

There are four statutory requirements for MDAP acquisition at Milestone A: the TDS, discussed in the previous section, a Consideration of Technology Issues, a Market Research report, and a CCA Compliance report. The Consideration of Technology issues is discussed in DoD 5000.2, and also in 10 United States Code (U.S.C.) 2634. Information about the Market Research report is available from 10 U.S.C. 3387 and 15 U.S.C. 644(e)(2). The CCA Compliance report is addressed in 40 U.S.C. Subtitle III Section 8088. [57]

In addition to the statutory information, there are eight regulatory information requirements for MDAPS Milestone A. Information about the specific requirements for all eight is available in DoD 5000.2. [230] The regulatory information required includes the ICD, the AoA, a Component Cost Analysis, a Cost Analysis Requirements Description, a Systems Engineering Plan, a Test and Evaluation Master Plan, Exit Criteria, and an Acquisition Decision Memorandum. [57]

2.3.3.4 Technology Development

The technology development phase for all MDAPs other than ships is still considered to occur before the initiation of a new acquisition program. The DoD has chosen to separate the technology development from the actual acquisition program in order to gain a more thorough understanding of the actual technology maturation. The purpose of this technology development phase is to allow necessary technologies to develop, under the guidance of the TDS and ICD. At the point where decision makers feel that the technologies have reached an acceptable level of military usefulness and have been proven in a relevant environment the Milestone B review is held. In most cases, because of the evolutionary nature of technology in most acquisition programs, the technology will not be fully developed at program initiation. [57]

In order to support program initiation, the targeted system user is responsible for developing a Capability Development Document (CCD). The CCD synthesizes information gained about the relevant technologies during the Technology Development phase and incorporates them into the context of the capabilities desired. This document replaces the ICD during later program phases. [57]

2.3.3.5 Milestone B

Milestone B contains a breakdown of statutory and regulatory requirements similar to Milestone A. It includes the statutory requirements of Milestone A, but adds the items shown in the first column of Table 4. The regulatory requirements of Milestone B include those of Milestone A and add those requirements shown in the second column of Table 4. These requirements outline the basic set of documents necessary to begin a MDAP. They do not reflect those required for a MAIS, or a MDAP-ship. The Milestone B is considered the formal initiation of the MDAP in most cases. [57]

Table 4: Additional Requirement for Milestone B

Statutory Requirements	**Regulatory Requirements**
Registration of Mission-Critical and Mission-Essential Information Systems	Acquisition Strategy
Benefit Analysis and Determination	System Threat Assessment
Programmatic Environment Safety and Occupational Health Evaluation (PESHE)	Technology Readiness Assessment
Spectrum Certification Compliance	Independent Technology Assessment
Selected Acquisition Report	Command, Control, Communications, Computers, and Intelligence Support Plan (C4ISP)
Live-Fire Waiver & Alternate LFT&E Plan	Affordability Assessment
Industrial Capabilities	Operational Test Agency Report of Operational Test and Evaluation Results
LRIP Quantities	Program Protection Plan
Independent Cost Estimate (CAIG) and Manpower Estimate	
Core Logistics Analysis/Source of Repair Analysis	
Competition Analysis	

2.3.3.6 System Development and Demonstration

DoD 5000.2 explicitly outlines "the purpose of the [System Development and Demonstration] SDD phase [as] development a system or an increment of capability; reduc[tion of] integration and manufacturing risk (technology risk reduction occurs during Technology Development); ensur[ance of] operational supportability with particular attention to reducing the logistics footprint; implement[ation of] human systems integration (HSI); design for producibility; ensur[ance of] affordability and the protection of critical program information (CPI) by implementing appropriate techniques such as anti-tamper; and demonstrate[ion of] system integration, interoperability, safety, and utility." [230] This phase is the non-technology related system design, and brings the system from a defined alternative to a producible system.

2.3.3.7 Milestone C

Milestone C is the final gateway before the system transitions into production and deployment. The statutory and regulatory requirements for milestone C are shown in Table 5. In each case, the documents must be updated to reflect the most current state of the program. Additional information about each requirement is available from [57].

Table 5: Milestone C Requirements [57]

Statutory Requirements	Regulatory Requirements
Consideration of Technology Issues	Initial Capabilities Document
CCA Compliance	Capability Production Document
Registration of mission-critical and mission-essential information systems	Acquisition Strategy
Benefit Analysis and Determination	Analysis of Alternatives
Programmatic Environment Safety and Occupational Health Evaluation (PESHE)	Systems Engineering Plan
Spectrum Certification Compliance	System Threat Assessment
Selected Acquisition Report	Technology Readiness Assessment
Industrial Capabilities	Independent Technology Assessment
Independent Cost Estimate (CAIG) and Manpower Estimate (reviewed by OUSD(P&R))	Command, Control, Communications, Computers, and Intelligence Support Plan (C4ISP)
Core Logistics Analysis/Source of Repair	Affordability Assessment
Competition Analysis	Component Cost Analysis
Technology Development Strategy	Cost Analysis Requirements Description
Acquisition Program Baseline	Test and Evaluation Master Plan
Cooperative Opportunities	Operational Test Agency Report of Operational Test and Evaluation Results
	Program Protection Plan
	Systems Engineering Plan
	Exit Criteria
	Acquisition Decision Memorandum

2.3.3.8 Final Phases

Production and Deployment and Operations and Support make up the final two phases of the acquisition process. The Production and Deployment phase assesses the operational effectiveness of the systems once obtained off the production line. In addition, it focuses on the development of the necessary capabilities for the actual manufacture of the system. [62] Operations and Support provides engineering support through the life-cycle of the system. [57] In cases where deficiencies in the field performance of the system are identified, an analysis is conducted to determine whether the loss in effectiveness warrants an update of the design.

The final phases of the defense acquisition system also require engineering analyses for the sustainment of the fielded system. This sustainment can include maintenance procedure updates, small scale component re-design, training, and end-of-service-life considerations. Planning must include the maintenance, supply, training and disposal of the system.

2.4 Analysis of Alternatives

Based on the “window of opportunity” concept presented by Ford and Barkan, it is undesirable to address robustness when only a small amount of design freedom is available. Also, as Mavris shows, selection of a design point locks down a great amount of design freedom. Therefore, because in most cases the Analysis of Alternatives is the DoD process which selects the single design and locks down the design freedom, the evaluation of robustness should occur during or prior to the Analysis of Alternatives. The following sections will explore the current guidance available from the DoD for the conduct of the Analysis of Alternatives.

AoAs are mandated by the DoD for all major acquisition programs, though the process for conducting an AoA is not explicitly directed by the DoD. The DAU is the primary source for DoD guidance to all branches of the military with regard to the acquisition process, of which the AoA is a part. The timing of the AoA in the overall defense acquisition process is shown by the light blue box in Figure 11.

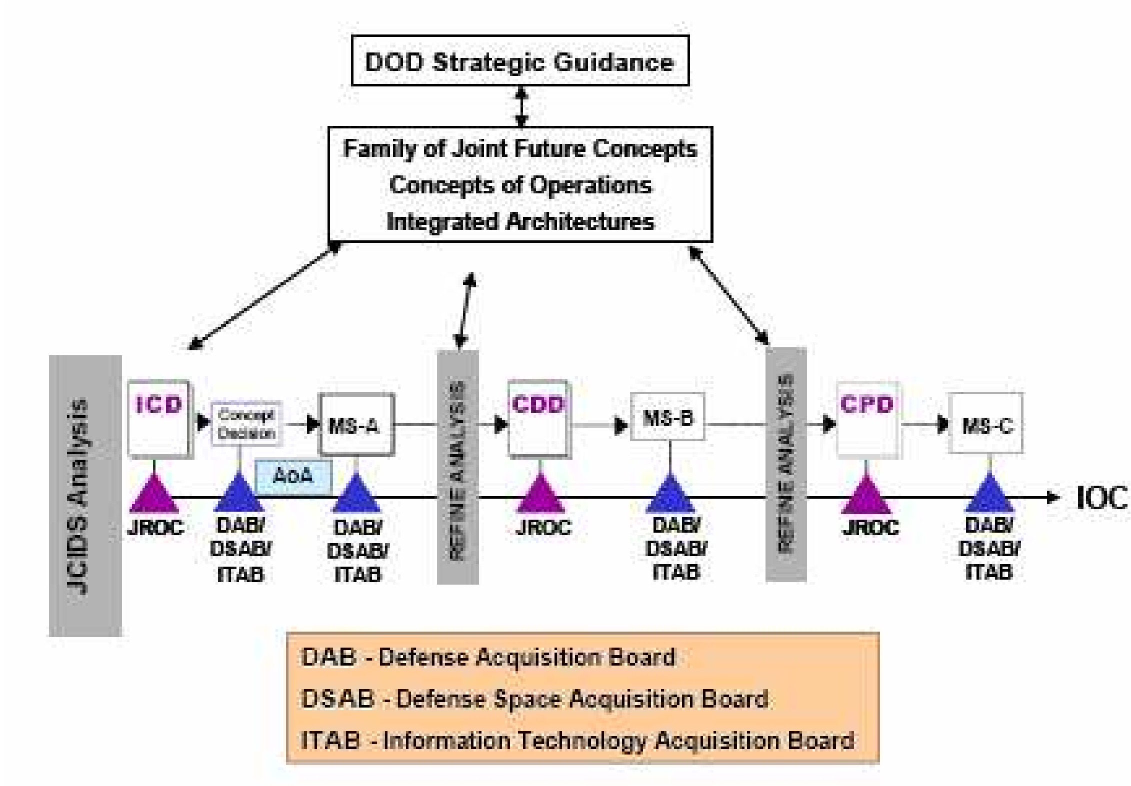


Figure 11: Defense Acquisition Process [203]

Each service maintains its own guidelines for the conduct of an AoA, within the framework set forth by the DAU. These guidelines are more specific than those published by the DAU, but vary in scope among the services.

2.4.1 Definition and Directive

According to the DoD, an Analysis of Alternatives is defined as "the evaluation of the performance, operational effectiveness, operational suitability and estimated costs of alternative systems to meet a mission capability. The AoA assesses the advantages and disadvantages of alternatives being considered to satisfy capabilities, including the sensitivity of each alternative to possible changes in key assumptions or variables. The AoA is one of the key inputs to defining the system capabilities in the capability development document." [203] The DoD specifies that the AoA is a mandatory procedure for MDAPS and MAIS Acquisition Programs. [230]

The DoD, while not specifically outlining the process or steps involved in conducting an AoA, does provide guidance with respect to necessary components and goals of the AoA. The DoD specifies that the AoA shall assess multiple elements of project or program alternatives including "technical risk and maturity, and cost." [230]

> The analysis shall be quantitative, and induce decision makers and staffs at all levels to engage in qualitative discussions of key assumptions and variables, develop better program understanding, and foster joint ownership of the program and program decisions. There shall be a clear linkage between the analysis of alternatives, system requirements, and T&E MOEs [Test & Evaluation Measures of Effectiveness] (Pub.L.104-106 (1996), Section 5123 and 44 U.S.C.3506). The analysis shall reveal insights into the program knowns and unknowns, and highlight relative advantages and disadvantages of the alternatives being considered. The activity conducting the analysis shall document its findings. [230]

The quantitative AoA should allow personnel involved with a project to make transparent decisions regarding the selection of system alternatives. By discussing and documenting key assumptions and variables, the thought process for discarding or further developing a particular option can be understood by later project reviewers. Additionally, the AoA may help identify potential problem areas that could emerge as the program progresses.

> The analysis shall include sensitivity analyses to possible changes in key assumptions (e.g., threat) or variables (e.g., selected performance capabilities). The analysis shall explicitly consider continued operating and support costs of the baseline. Where appropriate, the analysis shall address the interoperability and commonality of components or systems that are similar in function to other DoD Component programs or Allied programs (see 10 U.S.C.2457). For each alternative, the analysis of alternatives shall consider requirements for a new or modified

> [Information Technology] IT, including a [National Security System] NSS, or support infrastructure. [230]

The use of sensitivity analyses allows the evaluators to assess how well the system will perform in off-design conditions. The off-design performance is particularly important in the realm of military system design, as true operational conditions are difficult, if not nearly impossible, to predict. Battlefields evolve and new threats emerge that are often unanticipated by military planners. Additionally, as warfare becomes more asymmetric, the adaptability of the enemy becomes a large factor in uncertainty around system operating conditions. By varying the key assumptions of a system and observing the sensitivity, the off-design performance may be gauged, and the true affordability of the system understood. In this definition, affordability is precisely the ratio of the system performance to the life cycle cost of the system.

> The analysis shall aid decision-makers in judging whether any of the proposed alternatives to an existing system offers sufficient military and/or economic benefit to justify the cost. For most systems, the analysis shall consider and baseline against the system(s) that the acquisition program will replace, if they exist. The analysis shall consider the benefits and detriments, if any, of accelerated and delayed introduction of military capabilities, including the effect on life-cycle costs. PA&E [Program Analysis and Execution] shall assess the analysis of alternatives in terms of its comprehensiveness, objectivity, and compliance with the Clinger-Cohen Act... PA&E shall provide the assessment to the DoD Component head or Principal Staff Assistant (PSA), and to the MDA. The PM and MDA shall consider the analysis, the PA&E assessment, and ensuing documentation at Milestone B (or C, if there is no Milestone B) for ACAT I and IA programs. [230]

The AoA, if deemed to be acceptable in terms of methodology soundness and objectivity, serves as the primary decision making tool in the early phases of the defense acquisition process. Given the very large financial outlay of any MDAP, the decision makers attempt to focus objectively on the expected impact of the system, once obtained. This impact could be an increase in the capability of the military as a result of the system, or a maintained level of capability for a reduced cost. It is unlikely that military planners would accept a system from the AoA that reduced military effectiveness.

> Coordination shall ensure consideration of the full range of alternatives; the development of organizational and operational plans, with inputs from the Commanders in Chief of the Combatant Commands, that are consistent with U.S. military strategy; and the consideration of joint-service issues, such as interoperability, security, and common use. USD(AT&L) [Undersecretary of Defense (Acquisition, Technology, and Logistics)] shall issue guidance for ACAT ID programs. USD(AT&L) or ASD(C3I) [Assistant Secretary of Defense, Command, Control, Communications, and Intelligence] shall issue guidance for other programs. The Director, PA&E shall prepare the guidance in coordination with the offices listed above. [230]

In order to avoid the tendency of services to automatically "go with what they know" instead of considering the full range of alternatives, special attention must be paid to properly populating the alternatives space for the program. This should include not only system alternatives, but process alternatives with regard to the entire program life cycle including Research, Development, Testing, and Evaluation [RDT&E], manufacturing and operation. A great deal of complexity is added to the problem when assessing the full combinatorial range of alternatives, from a computational workload standpoint. Providing traceability through a design space of a million or billion possible alternatives also poses a challenge to the AoA team.

For the actual conduct of the AoA, it the DoD has left the decision making to the individual services or appropriate program managers. According to the USAF Office of Aerospace Studies (OAS) DoD 5000.2-R assigns the responsibility for preparation of the AoA to the service responsible for the mission area for which the capability need is determined. [172] The OAS provides extensive documentation on the conduct of an AoA, and offers an educational program. The United States Navy (USN) Office of the Assistant Secretary of the Navy has published some guidelines on the conduct of an AoA, but not nearly to the depth of the USAF literature. There appears to be no available United States Army (USA) documentation on AoA's available to the public.

2.4.2 USAF AoA Process

The USAF OAS at Kirtland AFB, New Mexico, offers three forms of educational material about AoAs. The first is the USAF Analysis Handbook, a 125 page document that outlines in detail the USAF standard process for conducting an AoA. Additionally the OAS offers a web-based short course for the conduct of AoA's and a live instruction in two possible formats: a course taught at Kirtland AFB, or an instructor sent to the unit involved in the conduct of the AoA. At this time, it is not clear if government contractors can participate in the short course options, or if it is limited to military personnel and government employees.

2.4.2.1 USAF AoA Format

OAS provides basic guidance on how to conduct each phase of the AoA based on the outline shown below. The guidance from OAS includes who is responsible for each section of the work, what should be done in each section, but not necessarily appropriate tools for each section.

Table 6: Organization of USAF AoA Report [173]

1. Introduction
   1. Background
   2. Purpose
   3. Scope
2. Acquisition Issues
   1. Mission Need
   2. Scenarios
   3. Threats
   4. Environment
   5. Constraints and Assumptions
3. Alternatives
   1. Description of Alternatives
   2. Nonviable Alternatives
   3. Operations Concepts
4. Determination of Effectiveness Measures
   1. Mission Tasks
   2. Measures of Effectiveness
   3. Measures of Performance
   4. Effectiveness Analysis
      a. Effectiveness Methodology
      b. Models, Simulations, and Data
      c. Effectiveness Sensitivity Analysis
5. Cost Analysis
   1. Life Cycle Cost Methodology
   2. Models and Data
   3. Cost Risk Methodology
6. Cost-Effectiveness Comparisons
   1. Cost-Effectiveness Methodology and Presentations
   2. Cost-Effectiveness Criteria for Screening Alternatives
7. Organization and Management
   1. Study Team/Organization
   2. AoA Review Process
   3. Schedule

   A. Acronyms
   B. References
   C. Lessons Learned
   D. Other Appendices

Throughout the AoA process, the OAS emphasizes the need for capability based analysis. Specifically, the goal is to determine how each alternative contributes to, or detracts from, the overall military mission accomplishment capability and the cost for that capability. From this standpoint, OAS has adopted the DoD emphasis on capability based design and decision making.

In addition to emphasizing capability-based analysis, OAS highly encourages the use of quantitative methods wherever possible in order to promote traceability. The traceability provides reviewing officers, as well as other personnel not present for the entire AoA process, a faster catch-up process, and allows decision makers to more fully understand prior decisions and the impact of their decisions. [173] A more rigorous treatment of qualitative decisions would aid the AoA process, and allow the traceability to extend from the quantitative analyses to the qualitative and overall system evaluation as well.

Figure 12 shows additional OAS suggested references for AoA's and related acquisition concepts.

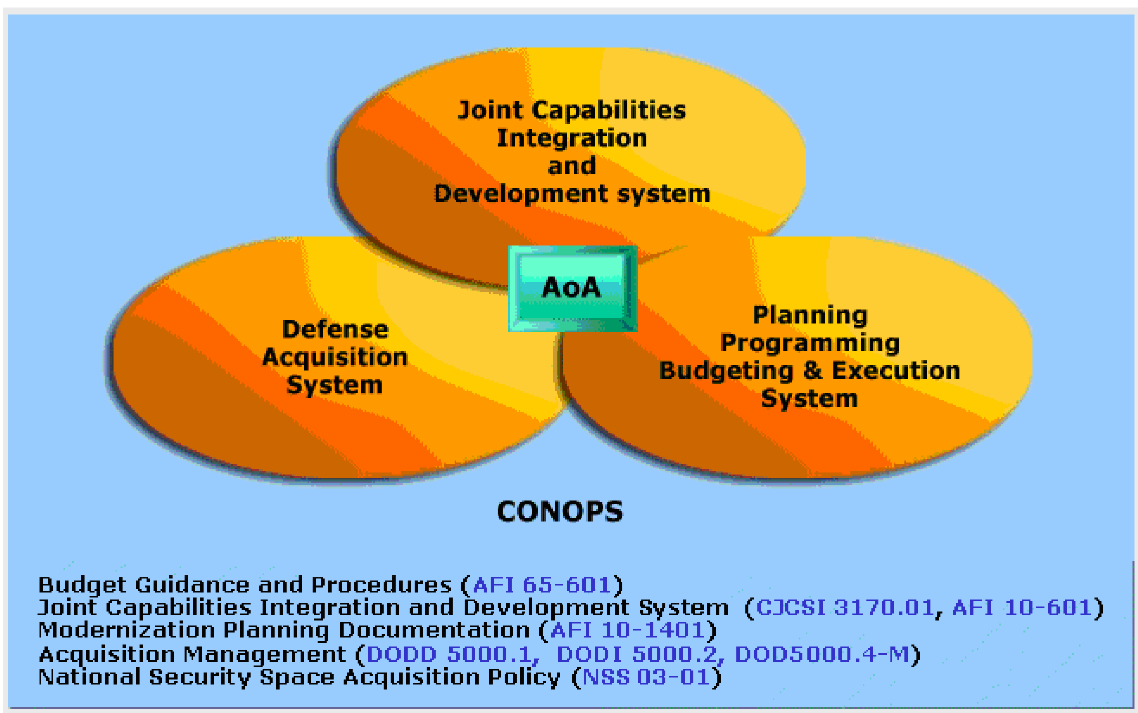


Figure 12: USAF AoA References [173]

2.4.3 USN AoA Guidelines

The Office of the Assistant Secretary of the Navy released guidelines for the conduct of Navy AoA's because of the fact that "DoD 5000.2-R places the responsibility for preparation of the AoA clearly on the organizational entity responsible for the mission area in which the requirement is determined to exist." [6] According to the documentation released by the Navy, the goal of an AoA is to determine if the best approach to meet the threat with respect to performance and resources expended.

The key areas identified by the Navy for an AoA are:

- Mission Need, Deficiencies and Opportunities
- Threats
- Operational Environments
- Operational Concept
- Alternatives
- Measures of Effectiveness (MOEs)
- Life-Cycle Costs of each alternative
- AoA (i.e., the actual analysis) [6]

The first four bullets above correspond roughly to the "Acquisition Issues" section of the OAS approach to AoA. The fifth and sixth bullets are approximately one to one with "Alternatives" and "Determination of Effectiveness Measures," respectively, in the OAS document. The seventh bullet roughly corresponds to the "Cost Analysis" section used in the USAF programs; however, the final "AoA" bullet appears to refer to sections that would be included throughout the OAS outline. Thus the Navy's direction for the conduct of an AoA includes similar information to the USAF process, but organized in a different fashion. The Navy does not, however, direct that "Cost-Effectiveness" be used as the primary analysis for the comparison of concepts.

The Navy breaks the AoA into sections for each milestone review in the program breakdown. The three milestone system follows a general program flow with more refinement on analyses as the program progresses. It focuses the AoA as a tool for program evaluation by decision makers at each milestone review.

> At MS [Milestone] I the analysis focuses on broad tradeoffs available between a large range of different concepts. The analysis normally presents a "Go/No Go" recommendation. It demonstrates why a new system is better than upgrading/modifying an existing system. Cost estimates may be only a rough order of magnitude but, nevertheless, an estimate is required. MS I AoA helps the MDA choose a preferred system concept and decide whether the cost and performance of the concept warrants initiating an acquisition program. MS I AoA can also illuminate the concept's cost and performance drivers and key trade-off opportunities; and provides the basis for the establishment of operational performance threshold and objective values for use in the ORD, APB, and Test and Evaluation Master Plan (TEMP). [6]

The wording above indicates that the AoA in the Navy is focused on the selection of a new alternative, with the current system or an upgrade viewed as a baseline. However, in many cases the upgrade of a current system is in fact the most cost-effective way to achieve a desired capability. For the AoA to be truly capability driven, the AoA should consider upgrades as equal alternatives with new systems.

> At MS II the analysis would be more focused. Hardware alternatives present a narrower range of choices. The analysis is more detailed than at MS I and more defined cost data are available. Point estimates are given with uncertainty ranges. Life cycle costs are normally presented.

> At production approval (MS III) the AoA, if required, is normally an update of the MS II document. It highlights any trade-off or cost changes. However, since cost and performance issues have typically been resolved prior to MS III, an AoA is not often required to support this MS. [6]

In essence, the Navy's guidance is recommending an AoA that increases in fidelity as the program progresses. However, the focus of the analyses are still on selecting a point solution, which does not lend itself well to updating of information based on new knowledge gained in the design process.

The Navy also specifies roles for the oversight of the AoA, the Analysis Director (who shall be independent of the program manager), the CNO/Sponsor, and the Program Manager. The role of an AoA in relation to multi-disciplinary analysis is mentioned, though it is not fully explained. According to the documentation, the AoA should progress as follows:

- Planning.
- Determination of performance drivers.
- Determination of cost drivers.
- Resolution of cost/performance issues.
- Preparing final briefing, and final report, if necessary [6]

A flow chart of the Navy's AoA process is included as Figure 13.

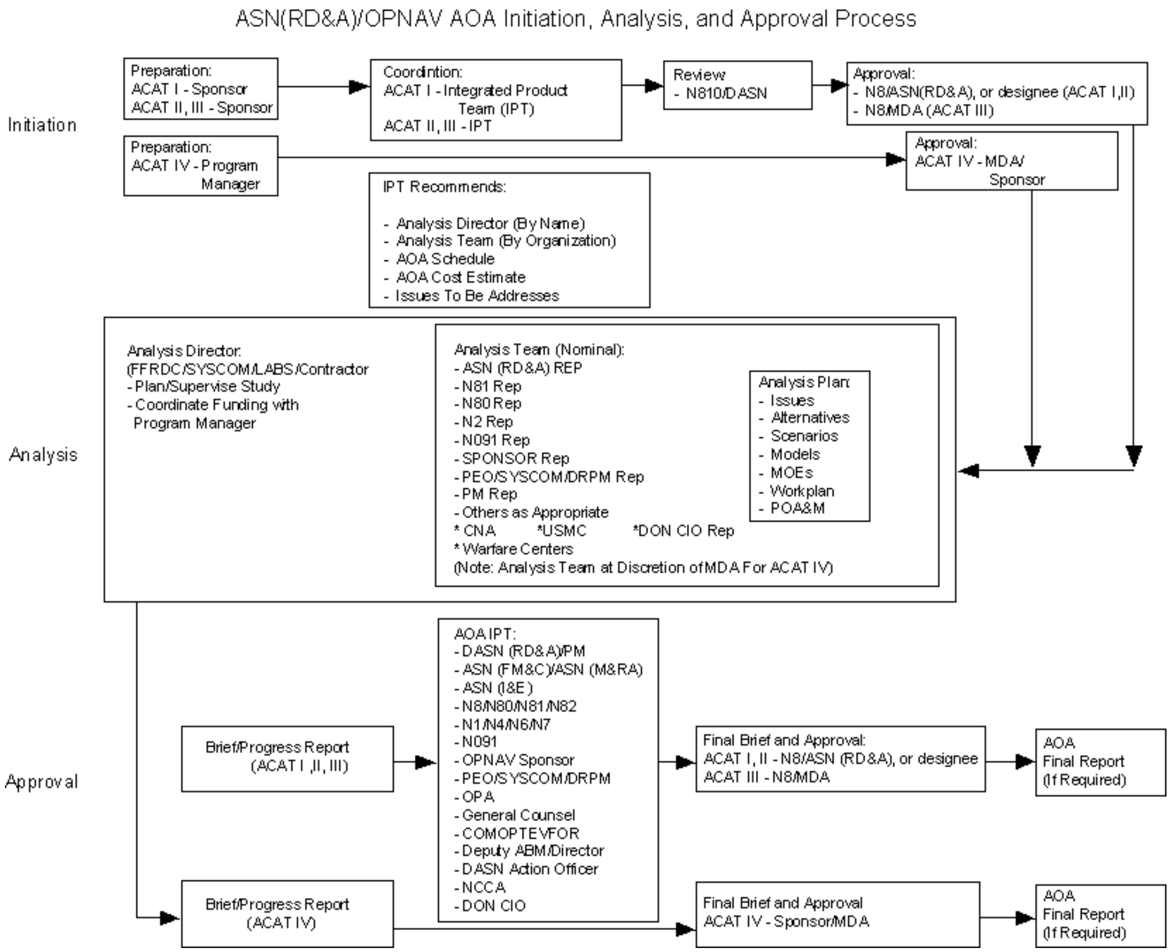


Figure 13: USN AoA Process Flowchart [6]

2.4.4 AoA Guidance Discussion

Both the USAF and USN have published fairly in-depth information with regards to the conduct of the AoA. In the case of the USAF, the emphasis is on a quantitative process that results in a point solution for further development in the defense acquisition process. In the case of the USN, the cost-effectiveness is not emphasized, but the AoA is revisited in at each milestone of the acquisition process. In the USN approach, the development of the AoA is very much like the spiral development of military systems, where the general system comes online, and is then upgraded as technologies mature and subsystem capabilities increase.

There is some difficulty with the assignment of the point solution in both the USN and USAF AoA processes. Updating the analysis as the program proceeds and information about technology maturation, policy issues relating to project funding, and future operating conditions can be very tedious once a design has been select. Essentially, the entire analysis must be conducted again with a new set of assumptions that better reflect the true development of the technology, political, and operating conditions.

2.4.5 Baseline AoA: KC-135 Recapitalization

Current US doctrine for the conduct of war delays major ground operations until air superiority is established. This approach has many benefits, but, perhaps most importantly, it delivers heavy damage to the enemy while exposing US personnel to minimum risk. Key to the establishment of air superiority is the effectiveness of air-to-air refueling systems, which greatly extend the range and endurance of air assets. The extension of sorties is particularly important early in air-superiority operations when friendly airfields may be sparse or non-existent. Aerial refueling tankers also allow strikes to originate from the continental US, and have the ability to keep surveillance aircraft aloft limited only by crew endurance.

The US refueling mission is primarily carried out by the KC-135E aircraft. These aircraft were originally commissioned in 1957 and, like the US Army's helicopter fleet, are becoming increasingly costly to operate. It was decided, therefore, that the KC-135 fleet should be recapitalized through upgrades or new acquisitions to allow cost-effective attainment of air power goals. A recapitalization is specifically defined as "The rebuild and selected upgrade of currently fielded systems to ensure operational readiness and a zero time, zero mile system. The objectives include: (1) extend service life; (2) reduce operating and support costs; (3) improve reliability, maintainability, safety, and efficiency; (4) enhance capability; and (5) reduce footprint on the battlefield. [218]

The initial plan for KC-135 Recapitalization was presented to the Congress; however, it came under extreme scrutiny due to illegal contract negotiations and lack of an Analysis of Alternatives. Senator John McCain, a member of the Senate Armed Services Committee, requested an AoA as required for all defense acquisitions of this scale, and consistent with 5000.2-R. [187] The AoA was to be conducted by a Federally Funded Research and Development Center (FFRDC) or other independent agency. The RAND Corporation's Project Air Force (PAF) was selected to conduct the AoA, with the Institute for Defense Analysis (IDA) checking soundness of methodology and objectiveness. The purpose of the AoA was to ensure that the most cost effective alternative for recapitalization of the KC-135 fleet was selected. [135][134]

The alternatives for the recapitalization study were provided to the RAND PAF by the Acting Under Secretary of Defense, Acquisition, Technology and Logistics, Michael W. Wynne. The set of alternatives is shown below in Table 7, and does provide a good variety of aircraft for consideration across a number of aircraft types. In addition, fleets consisting of combinations of the alternatives in the table were considered in the AoA. [135]

Table 7: KC-135 Recapitalization Alternatives

Category	**Alternatives**
New, Commercial Derivative Tankers	Airbus 321, 330,340,380 Boeing 737,767,787,777,747
Used Commercial Derivative Tankers	Airbus 310, 330 Boeing 757, 767, 747 DC-10, MD-11
New Military Derivative Tankers	C-130J, A400M, C-17
Newly Designed Tankers	Boeing, Lockheed Martin, Northrop Grumman, Aeronautical Systems Center
Newly Designed Tanker Transports	Unnamed (5)
Unmanned Aerial Vehicles (UAVs)	Unnamed
Stealthy Tankers	Unnamed (2)
Commercial Sources	Unnamed

The RAND PAF presented the following questions as the focus of the AoA, which will be discussed in the baseline discussion section

KC-135 Recapitalization Research Questions:

1. What is the most cost-effective alternative for recapitalizing the KC-135 fleet? (Here, an "alternative" can be a fleet consisting of a single type of aircraft or a fleet consisting of more than one type.) Again, in this AoA, the most "cost-effective" alternative means precisely the alternative whose effectiveness meets the aerial refueling requirement at the lowest cost.

2. When should the recapitalization assets be acquired?

In addition to the cost effectiveness and recapitalization timelines, the AoA considered two additional criteria for the recapitalization assets: operational concerns in terms of airfield use, and versatility in terms of cargo and passenger capacity. Both of these areas were considered for each alternative in the AoA, but their impact on the selection of a concept was considered a "matter for senior decision maker judgment." [135]

With the research questions in place, the groundwork was complete for the comparison of the alternatives in the alternatives set. The alternatives were compared using the approach of fixing the effectiveness of each type of fleet, and then comparing the complete life-cycle cost necessary to achieve that level of effectiveness with the aircraft. [187] A summary of the RAND PAF methodology for comparing the alternatives, defined by the USD(AT&L), is included in Figure 14.

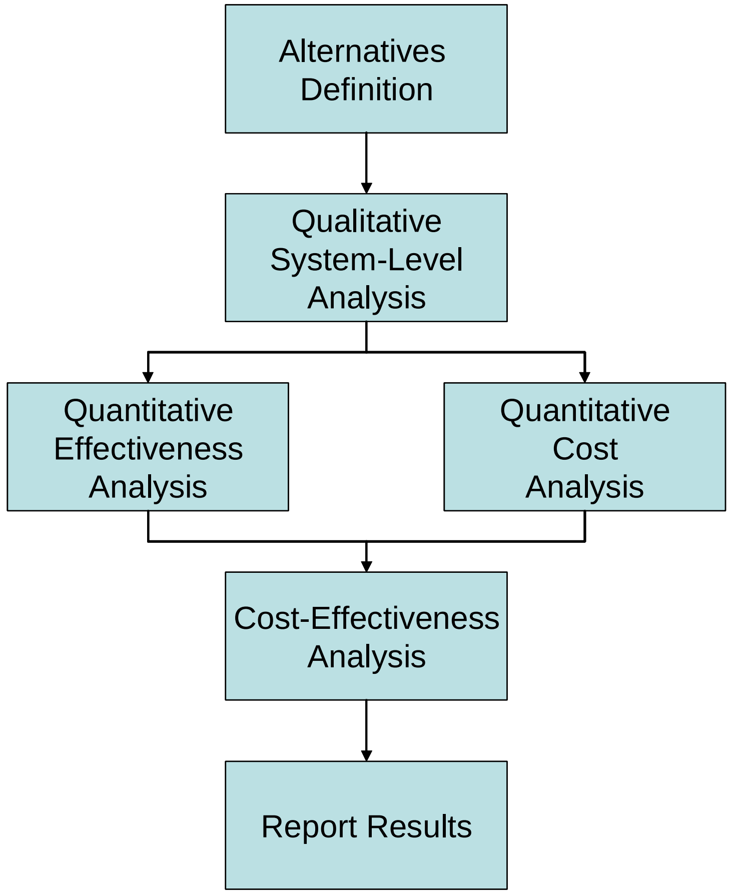


Figure 14: RAND PAF AoA Methodology [134]

The set of alternatives was compared for a variety of future operating conditions, including, refueling requirement, operational characteristics of the refueling aircraft, technical performance of the tankers, the configuration of the tankers, differing cost projections, and the planning horizon for the analysis. [135] While RAND PAF reports that there were a wide range of possible future operating conditions considered, in the publically available documentation there is no reference to how many cases were considered, what the ranges on the variables defining the operating conditions were, or how much the operating condition impacted the results of the study. The only comment RAND PAF makes with regard to the sensitivity of the results of the AoA to future conditions is to say that "the results hold true regardless of the specific projection of the factors within the broad ranges examined." [135]

The RAND PAF recommended that a fleet of medium to large commercial derivative tankers be acquired for the recapitalization of the KC-135, as they were the most cost-effective alternative. [187] The decision for timing of the recapitalization is not driven by the cost-effectiveness metrics considered for this study and therefore should be made based on other factors of interest to the DoD. [135]

2.4.6 KC-135 Recapitalization Discussion

Because the metric of primary concern in the recapitalization of the aerial refueling tanker fleet is the effectiveness and cost of that fleet, the problem must be viewed as a system-of-systems. The effectiveness of the fleet with include the size, operation and architecture of the fleet itself (the system-of-systems), the characteristics of the aircraft conducting the missions (the systems), and the characteristics of the crew, fuel volume, etc within each aircraft (the sub-systems). In the RAND study there were several allusions to considerations of costs at all system levels, but the alternatives that were defined by the Under Secretary of Defense were only systems. By limiting the conceptual design space in this fashion, the Under Secretary removed two levels of the system-of-systems hierarchy, and consequently limited the possible effectiveness of the Analysis of Alternatives itself. The interactions between the levels of a system-of-systems often limit the capability of that system, and without considering the entire synthesized system-of-systems, those interactions are ignored.

While identifying the most cost-effective alternative as defined above seems like, at first glance, a logical way to compare candidate alternatives, it does not address the system of systems approach necessary for truly evaluating the merits of a complex system. By locking the requirements in place for the refueling fleet, the critical dimension of the impact of evolving requirements is ignored. While changing requirements were addressed in the form of alternate scenarios to check the robustness of the system to a number of different mission requirements, there is no evidence that the requirements were treated as independent variables so their impact on the system could be studied in detail.

Because the refueling fleet is a system within a larger system of systems, a capability based approach should be adopted as it allows the AoA to incorporate changes at many system levels that could result in a more effective or least costly system. Rather than rely on "cost-effectiveness" as defined above, the analysis should be conducted by analyzing the "affordability" of a system solution. The precise definition of affordability is the ratio of the performance of a system to the cost of achieving that performance.

It should again be reiterated that this baseline study was conducted on the UNCLASSIFIED version of the summary report, as well as unclassified presentations made available by the RAND PAF. It is possible that in the SECRET version of the report, a different set of alternatives, scenarios, etc were explored. In the absence of this information however, the assumption will be that the summaries available were representative of the entire effort.

2.5 Summary

The important terms relating to the title of the dissertation were defined and discussed with preference for DoD definitions wherever possible. Robustness was defined as the ability of a system to perform over a wide range of conditions. Evaluation is the determination of worth (usually the ratio of benefits to costs), through careful appraisal and study. A SoS is a set or arrangement of systems that are related or connected to provide a given capability. Regret is the difference between a system's evaluation metric(s) and the best system's evaluation metric(s) for a scenario. And finally, analysis is the systematic examination of a complex by considering its elements and their relations.

A brief overview of the activities, phases, and tasks of the defense acquisition process was given. The color coded rows of the Integrated Defense Acquisition, Technology and Logistics Life Cycle Management Framework correspond to the JCIDS process (high-level military planners), the Defense Acquisition System (acquisition specialists and industry), and the PPBE (government financing). The 6 phases of the process are coded by the color of the task boxes and are divided by milestones or decision markers.

Based on the "widow of opportunity" concept from Robust Concept Design, the early phases of the defense acquisition process offer the most potential impact for improvement of the SoS products and processes. A more detailed discussion of the AoA activity in the Pre-Milestone A Defense Acquisition System is presented. Each service maintains its own procedures for the conduct of the AoA, though general guidance is passed from the Secretary of Defense and Undersecretary of Defense (AT&L) via the Defense Acquisition University.

The KC-135 Recapitalization AoA, performed by the RAND corporation, was examined as a baseline for the current state-of-the-art in AoAs. In the RAND study there were several allusions to considerations of costs at all system levels, but the alternatives that were defined by the Under Secretary of Defense were only systems. Additionally, while RAND PAF reports that there were a wide range of possible future operating conditions considered, in the publically available documentation there is no reference to how many cases were considered, what the ranges on the variables defining the operating conditions were, or how much the operating condition impacted the results of the study. The lack of a systematic study of the robustness of the candidates provides an area for improvement in the AoA state-of-the-art.

CHAPTER 3

RESEARCH FORMULATION

This chapter presents the research formulation for attempting to improve the ability of design engineers and decision makers to understand the robustness of alternatives early in the defense acquisition process. The process of generating this research formulation was iterative, included many thought exercises, and involved an extensive search of literature in both the aerospace engineering realm and other fields. Because the iterative nature of the formulation is difficult to convey in text, which flows linearly, the final state from each primary area will be presented. The sections presented in this chapter include (1) the intent of the dissertation, (2) the perceived gaps in the state-of-the-art and the desired state, the challenges associated with those gaps, (3) a set of high level research questions related to the gaps, (4) a discussion of the genesis of the hypotheses for filling the gaps, and (5) additional research questions that were created at a lower level because of requirements of the proposed solutions.

3.1 Dissertation Intent

The motivation chapters of this dissertation identified the robustness of military systems-of-systems as the area of interest based on several logic experiments and case studies. The goal of this research is **to improve the current state-of-the-art in early defense acquisition processes through increasing the engineer's and decision maker's ability to compare the robustness of competing alternatives**.

3.2 Assertions

There are several assertions that form the logical backing for this research objective. The path taken and the decisions made represent one of many possible approaches to looking at robustness and the defense acquisition process.

3.2.1 Assertion 1 – Defense Acquisition

Improvements to Defense Acquisition Process could improve MDAP performance

The motivation behind this dissertation was the ineffectiveness of current military systems in the current operational scenario and the cancellation of several MDAPs because of anticipated shortcomings, performance, situational appropriateness, or affordability. Anecdotal evidence from numerous people involved in the Defense Acquisition Process indicates that improvements are needed. It has also been observed that improvements in the development process for products in general usually result in improved products [215]. This logic is being extended to military systems. However, because it is not possible to test this assertion by designing two military systems for a control and test case, this assertion will be accepted without further attempt of proof. The acceptance of this assertion leads to the next assertion about when the improvements should be focused.

3.2.1.1 Assertion 1.1

Pre-Milestone A offers great opportunity for impact

The defense acquisition process is long and cumbersome. There are hundreds of possible tasks that could be improved upon. However, because the cost of change is lowest in the earliest phases of a design and the potential impact of change is greatest [158], the early phases should be focused on for improvement first. The Pre-Milestone A processes, specifically the JCIDS process and the AoA offer great potential for improvement, especially considering the impact the JCIDS process and AoA have on all other activities in the Defense Acquisition System. The National Research Council, working under a request from the Deputy Assistant Secretary of the Air Force, identified Pre-Milestone A as an area for improved systems engineering and noted that “about three-quarters of total system life cycle costs are influenced by decisions made before… Milestone A.” [11]

There are two schools of thought surrounding improvements in the early acquisition process phases. The first argues that decision freedom should be preserved for as long as possible to allow uncertainty to clear. The second argues that a decision should be made early on, but based on as much information as possible. Because these are two fundamentally different approaches to addressing the problem, but both are used in defense acquisition, the method will not be specifically tailored to either. Rather, every effort will be made to allow decision makers to use either philosophy while working with the method developed in this dissertation.

3.2.2 Assertion 2 – Robustness

Using robustness as a criterion for selecting among alternatives will improve SoS performance

Military systems are used across a wide range of scenarios, many of which may have never been considered when the system was first designed. The B-1 bomber was designed solely as a nuclear strike aircraft, but has become a conventional bomber with the dissolution of the Soviet Union and the development of the Joint Direct Attack Munition. [139], [241] The system received a new lease on life because it was robust to the change in the tactical environment in which the US operates. The unarmored HUMVEE has been upgraded with armored sides in order to increase its effectiveness in operating under conditions in Iraq that were unanticipated during its design. However, these upgrades are greatly increasing the engine wear on the vehicles and increasing their cost to operate. These are just two examples of current military systems that are operating in scenarios outside those included in their initial design, with varying degrees of success.

By definition, optimal systems will perform better than robust systems in the conditions for which they were designed. However, as noted by Borer, military operations are almost never at "on-design conditions," [28] the selection of a system that is robust will improve overall performance.

3.3 Gaps: Current State-of-the-Art and Desired State

The following gaps have been identified based on the case studies, evaluation of current robustness methods, and dissection of the defense acquisition process.

3.3.1 Gap 1

As was observed in the baseline study of the KC-135 recapitalization, the robustness of candidate alternatives is currently studied through a limited number of off-design simulations. This does allow a limited understanding of the robustness of a particular candidate, but not the robustness relative to the other candidates. Therefore the first gap is the **lack of a quantifiable metric for the robustness of a system**.

3.3.2 Gap 2

An additional problem with the current approach to assessing robustness of candidate alternatives is the limited nature of the off-design explorations that can be accomplished. This very limited nature is in stark contrast to the limitless number of ways that operational scenarios and enemy technologies can evolve. The second gap follows as the **inability to account for a massive possible scenario space in assessing robustness**.

3.3.3 Gap 3

The third gap in the current state-of-the-art and the desired state is the **difficulty in updating the Pre-Milestone A activities as additional information becomes available about future operating conditions and technology maturation**. The extended timeframe of MDAP development, decades in many cases, means that the knowledge of the operating conditions, while fuzzy at first, will become clearer as the program progresses. This is analogous to a cloud of uncertainty “shrinking” to a smaller cloud or a point as the program develops. Understanding the impact of the scenario maturation currently required a nearly complete rework of the Pre-Milestone A activities.

3.4 Research Questions

The following research questions were developed based on the gaps outlined in the preceding section and the research objective expressed in the dissertation intent section. The development of the research questions was an iterative process that included a thorough literature search of the aerospace engineering literature, the defense acquisition literature, and literature from other disciplines. The two research questions presented here provide the overall questions the research is attempting to answer; however, many other questions were considered in the process of addressing these.

3.4.1 Research Question 1

- Most Major Defense Acquisition Programs stretch for more than a decade, so how can we evaluate the robustness of candidate system-of-systems solutions while considering the uncertainty associated with:
  - Technology maturation?
  - Possible warfare doctrine?
  - Possible enemy set?
- How can we define robustness to include these uncertainties?

MDAPs naturally fall into an undesirable region of high uncertainty, because numerous assumptions must be made early in the defense acquisition process, and high risk, because a large amount of taxpayer dollars required to develop this class of SoS. The successful development of a complex SoS to a very high performance level with a long period of program development is inherently difficult. The development program is impacted by uncertainty with respect to many aspects of the SoS.

3.4.1.1 Cost and Performance Uncertainty

Uncertainty associated with the SoS cost and performance primarily relates to the accuracy of modeling and simulation techniques available for use at the early phases of conceptual SoS design and to the accuracy of assumptions that were made in the modeling and simulation variables. Typically, the models used in early SoS design contain a fair amount of uncertainty because of the speed at which they must be capable of evaluating SoS concepts. Additionally, there are assumptions about the performance of immature technologies that will be used in the final SoS. Often, the only information about these technologies is available from the researchers developing the particular technology, who are quick to sing praises but often hesitant to share the costs or problems with a new customer. Many revolutionary projects for the military rely heavily on emerging technologies to step ahead of current and potential adversaries. When so much money is involved in the development of a SoS, the SoS designer must be able to account for the possibility of a different maturation result for critical technologies.

3.4.1.2 Operational Environment Uncertainty

Uncertainty about the operating environment for the SoS appears because the possibility of differences in the assumed operational doctrine of friendly commanders and forces, and the possibility of differences in the set of enemies for which the SoS was designed. Every new war brings a new set of challenges, many unanticipated by planners. In the 20^{th} century, warfare evolved from trench based to large army maneuvers to more asymmetric methods. [210] In recent years, the pace of the shift in enemy tactics, and as a result the pace of US doctrine change, has greatly increased. Military planners of previous centuries saw evolutions in tactics that took many years to take hold. However, in regard to the current conflict, "[Former Marine Commandant] Hagee describe[s] Iraqi insurgents as clever fighters who change their battlefield tactics every seven to 10 days, making it difficult to stay ahead of them." [163] These rapid evolutions of enemy tactics mean that a point condition of the environment could have gone through 700-1000 evolutions during the duration of a 20 year MDAP. These rapid changes in operating conditions drive the need to greatly increase the robustness of military SoS.

3.4.1.3 Definition of Robustness

The current definitions of robustness outlined earlier in this dissertation do not easily allow the consideration of performance for many systems over a large number of operational scenarios. A mathematical definition that allows the assignment of a robustness metric to each candidate design and is a function of the design's performance and cost, as well as the relative likelihood of the scenarios under which it will be evaluated is desired. Additionally, a relatively simple definition is desired, both for clarity and for computational load while evaluating over large numbers of scenarios.

3.4.2 Research Question 2

- ❖ How can we promote the ability to update the robustness analysis as higher-fidelity information about the system-of-systems' operating conditions becomes available?

Information in early stages of systems design relies on assumptions, especially when dealing with immature technologies that are common in MDAPs. However, as technology and the SoS design mature, assumptions are replaced by more concrete information from modeling and simulation, bench tests, and finally full SoS field tests. As this information becomes available, however, it is rare to find the systems engineering tasks of Pre-Milestone A repeated. This is because of the cost and engineering time associated with improving upon the earlier analyses. However, important information about the SoS performance, especially in the context of a wide number of scenarios, can be gained by revisiting the early tasks.

The most apparent benefit would be increased understanding of the bounds of SoS performance across the possible operational scenarios. This information would allow better planning for future gaps in capability. Additionally, the updated early systems engineering studies would allow tweaking of design requirements where tasks remain unfinished. For example, if a radar technology matured to a lower-than-expected level, missile seeker requirements could be made more stringent to compensate for the loss of radar performance. Both of these benefits would not be realized without a cost-effective way to update the systems engineering analyses of the Pre-Milestone A period.

3.5 Hypothesis Genesis – Creating a New Method

3.5.1 Functions Required

The methodology of this section is proposed to fulfill a research objective: to incorporate robustness into the decision making process and encourage adaptability of the Pre-Milestone A period. There are eight generic tasks that must be completed in this phase [58].

1. Establish the need
2. Define the problem
3. Establish measures of performance (MoPs) and measures of effectiveness (MoEs)
4. Generate architectures
5. Generate alternatives
6. Analyze alternatives
7. Compare results
8. Make a decision

These tasks vary slightly from those that appear in some systems engineering literature, but the basic purpose of the tasks is the same. In many cases, systems engineering assumes that the first phase shown below is conducted by someone outside the organization doing the systems engineering. However, for defense acquisition the JCIDS process involves establishing the needs of the military. The generation of architectures and the generation of alternatives are separated because of the processes in the Pre-Milestone A phase of defense acquisition. Systems architectures are usually established before and guide the generation of alternatives. There is usually a down-selection among architectures before systems alternatives are created.

3.5.2 Scope

The methodology of this dissertation is not intended to address all of the activities in the Pre-Milestone A phase of defense acquisition. Such a task would be beyond the scope of a doctoral dissertation, and would require much coordination with the government agencies responsible for the defense acquisition activities. Rather, the methodology in this dissertation will only address those generic areas where the research objectives, questions, and hypotheses relate to the activities.

3.5.3 Activities Focus

The research objective for this dissertation most closely aligns with the analysis of alternatives and the comparison of results from the generic tasks list. There is additional impact in the establishment of MoPs and MoEs, in that new measures must be included.

- **Task 3: Establishing MOPs and MOEs**
- **Task 6: Analyze Alternatives**
- **Task 7: Compare Results**

3.5.4 Cross-Fertilization from Long-Term Policy Analysis

Because no existing systems engineering method for assessing robustness performs well enough for application to the early tasks of defense acquisition, other fields were searched for methods. The hope was that a method existing in another field, with support of experts and literature in that field, could be applied without modification, or with slight modification, to the defense acquisition process. There were several criteria identified by the author as an initial screening for finding appropriate fields for methods investigation: similar time-frame, existence of a large amount of uncertainty, and high-stakes/risk. After searching literature from a variety of disciplines, long-term policy analysis was identified as a promising field.

3.5.4.1 Time-Frame

Long-term policy analysis typically deals with a time horizon of somewhere between 10 and 50 years. [141] Military acquisition processes typically take between 5 and 20 years, though this fact has been lamented by military planners. [11] The similarity in these time-frames creates an environment where the following two filter criteria are more likely to be met.

3.5.4.2 Uncertainty

Lempert defines deep uncertainty as the condition when "when we do not know, and/or key parties to the decision do not agree on, the system model, prior probabilities, and/or "cost" function." [142] This is opposed to a system where the probabilities are well behaved, the system model exists and is readily understood, and the cost function is well defined. The more well-behaved case is close to Taleb's concept of "mild uncertainty" or "Gaussian uncertainty." [220] The concept of mild uncertainty is very applicable in near-term problems and corresponds to the majority of the methods that exist in systems engineering for evaluating robustness.

Military acquisition exists in the realm of deep uncertainty. The system model for military operations is poorly understood and rife with human-factors. Especially as the concept of network-centric warfare has come to dominate battlefield operations, simple statistical relationships, such as those established by Dupuy, [71] no longer are applicable. Additionally, the "cost function" for current military systems changes depending on the decision maker. Often there are unspoken constraints that drive designs and are never formally communicated to the designers.

3.5.4.3 Risk

The level of risk associated with long-term policy analysis is the same, or perhaps greater, than that of defense acquisition. James Dewar notes that the following are all examples of successful long-term policy [35].

- The US Constitution
- Panama Canal
- Transcontinental railroad in the US
- Marshall Plan
- Bismark's unification of Germany
- George Kennan's policy of "containment" of the USSR
- US Social Security plan
- FCC helping the US phone system connect to computers

All of the policies above, if unsuccessful, carried great potential consequences, ranging from a failed early United States to billions in economic losses to overseas competitors who could have adapted technologies before us.

While the consequences of failure are not as great for MDAPs as for these major policy decision outlined above, they are enormous. The loss of billions of dollars of taxpayer funds can derail political establishments and cause major corporations to fail. Because these two areas are in the same realm of risk, at least relative to most small risk calculations done in systems engineering, long-term policy analysis is an acceptable fit for identifying methods for cross-fertilization.

3.5.5 Analyzing Alternatives

The three primary areas where a contribution is being made will be discussed out of order. This is because the contribution for the establishment of MOE's and MOP's is a creation of this dissertation's author based on the contributions from long-term policy analysis to the other two primary areas. These two cross-fertilized ideas and their potential contribution to defense acquisitions are discussed in the following sections.

The first area of cross-fertilization, Massive Scenario Generation, allows the designer and decision maker to consider the utility of an alternative across a much wider set of possible future scenarios than was previously available. The second area of advancement, parametric methods, allows the designer to rapidly update analyses as information about the future becomes available.

3.5.5.1 Massive Scenario Generation

Massive Scenario Generation is an approach to exploring possible futures with the aid of computer models. The technique was developed at the RAND corporation for use in long-term policy analysis and for strategic planning. The development of this technique was dependent on the development of powerful computing capabilities that have recently become prevalent in the research environment.

Massive Scenario Generation was constructed to help humans consider the implications of policy decisions across a "very large landscape of plausible futures." [141] The ability of the policy decision to be implemented in a computer simulation that can realistically capture the dynamics of the problem is crucial to the validity of the Massive Scenario Generation results. In Lempert's formulation, Massive Scenario Generation is used to create "scenario ensembles," which are discrete cases intended to represent the landscape of plausible futures. Exploratory modeling software and a computerized scenario generator are used to construct the large set of scenarios that make up the scenario ensemble.

3.5.5.2 Defining a Scenario

A key part of defining a parametric scenario for Massive Scenario Generation is understanding how to categorize elements that belong to the scenario and identifying interactions among the elements that impact the alternative being evaluated. In Lempert's work on regret analysis coupled with Massive Scenario Generation, the RAND team used an extensive literature review to identify potential input variables and metrics, and then relied on the experts on the team to categorize and prioritize them. [141] For military alternatives analysis, the basic initial breakdown is suggested to be friendly systems (including the alternative being analyzed), targets, and the general environment. The general environment will include threats that are not targets, and the physical characteristics of the world. A sample breakdown is shown in Figure 15.

For the purpose of this dissertation, targets will be considered part of the environment. By considering the targets as a part of the environment, the scenario can be broken down into two groups: things over which the friendly side will have control and those things that they will not.

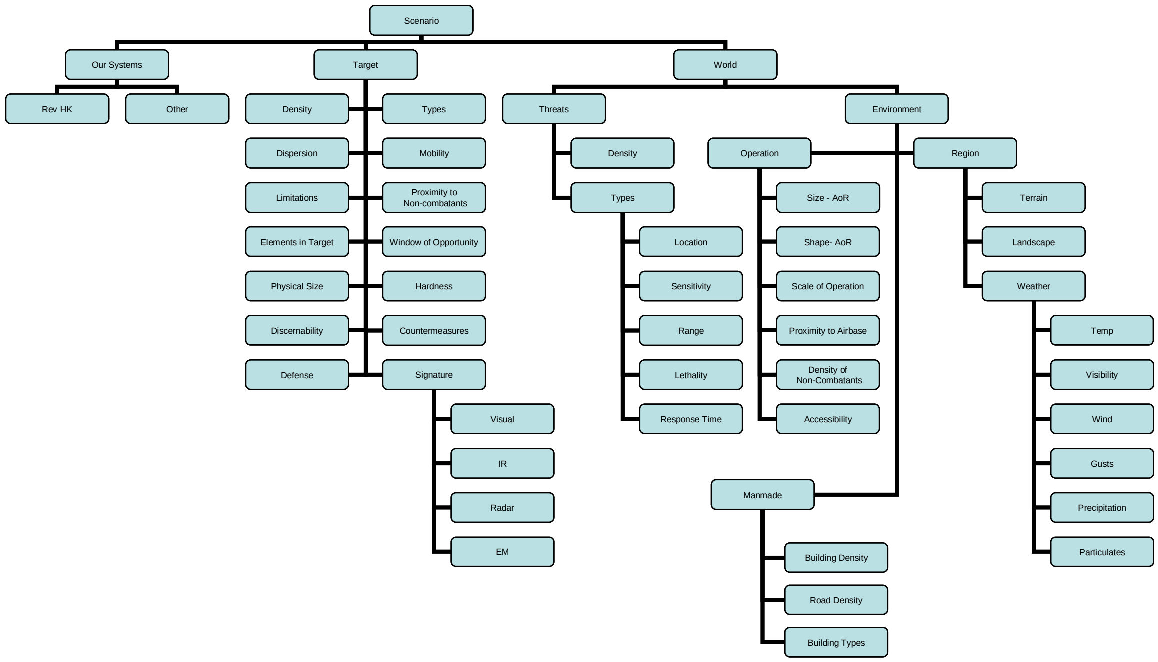


Figure 15: Scenario Breakdown

3.5.5.2.1 Environment

The principle concern for evaluating SoS concepts in a particular possible future is what makes up the possible future. The particular realization of events leads to an environment in which the SoS will function. The environment in which the SoS acts, combined with the actual matured state of the SoS itself, combine to form the future scenario.

The environment is defined in this dissertation as all of the factors which affect the SoS, but are not a part of it. This is based on the definition from Webster, which states that an environment is “the circumstances, objects, or conditions by which one is surrounded.” [4] The environment is made up of three subsets: the physical environment, the target environment, and the threat environment. However, relevance of each element of the environment subset, physical, threat, or target, will depend on the required level of fidelity for the problem at hand.

3.5.5.2.2 Target Environment

The SoS's target is "a geographic area, complex, or installation planned for capture or destruction by military forces." The intelligence community definition is "a country, area, installation, agency, or person against which intelligence operations are directed." For targeting purposes, this definition must be expanded to include the contents of the area, complex, or installation (e. g., people, equipment, and, resources). Furthermore, capture or destruction must be expanded to include "disruption, degradation, neutralization, and exploitation, commensurate with objectives and guidance." [236]

A target must qualify as a military objective before it can become a legitimate object of military attack. In this context, military objectives include those objects that, by their nature, location, purpose, or use, make an effective contribution to military action, or whose total or partial destruction, capture, or neutralization offers a definite military advantage. The key factor is whether the object contributes to the enemy's war fighting or war sustaining capability. Consequently, an identifiable military benefit or advantage should derive from the degradation, neutralization, destruction, capture, or disruption of the object. Not only does this concept preclude violations of the Law of Armed Conflict (LOAC), but it also supports the principles of war by employing economy of force against valid military objectives.

The target environment describes all aspects of the SoS's target that are relevant to the function or performance of the SoS. This definition is intended to include characteristics that may not intuitively be a part of the target itself, but nonetheless have an impact on the performance of the SoS. An example of this might include proximity of the site to a major religious site, which would limit the SoS's ability to apply energy to the target in many conflicts.

3.5.5.2.3 Threat Environment

The threat environment describes all elements of the adversary's assets that can potentially impact the SoS in an adverse way. These are outside the set of elements included in the target environment and can include 3rd party threats. The threat environment is considered to be entirely "man-made" and therefore, while some natural occurrences would be threatening to a SoS, forces of nature are not considered part of the threat environment.

3.5.5.2.4 Physical Environment

The physical environment will be defined as all elements of the environment that are not included in the target environment or the threat environment, but can affect the SoS or its performance with respect to the MOEs for the scenario.

3.5.5.3 Parametric Methods

Parametric methods, as opposed to deterministic methods, typically do not return a "single answer." Rather, a parametric method will focus on establishing a set of relationships that will return an answer for a range of input parameters. Input parameters correspond to the independent variables of a deterministic function (or method), but are allowed to take a range of values. [24]

Parametric methods have become important in the design of highly integrated systems, such as aircraft, because of the uncertain nature of many system aspects in early design phases. For example, historical data may be used in aircraft conceptual design to assign an anticipated weight to the aircraft's engines, upon which the structure design is dependant. If upon conducting a detailed design the engines are determined to be heavier than anticipated, the aircraft's structural weight will have to be increased, which either will require more thrust (an engine re-design) or will reduced performance. By using a parametric approach, however, designers can rapidly update the entire design by simply "dialing in" the new engine weight.

Baker's Unified Tradeoff Environment (UTE) [19] provides an environment in which parametric methods can be visualized through partial differential equations. Baker's formulation was initially implemented using a rotorcraft example, but has been extended to autogyros by Ahn [10], and to SoS by Biltgen and Ender [23], [25]. By viewing the partial differential equations in the UTE, not only can the designer "dial in" a new design and rapidly see the results, the impacts of the various parameters can be visualized simultaneously.

Parametric methods have an added benefit in the current paradigm of electronic design reviews. [156] Analyses presented to decision makers are rife with assumptions that have been made in order to enable the use of models, simplified relationships, and even many empirical tests. If a decision maker disagrees with an assumption, the entire study can be discredited in his or her eyes. However, if a parametric study is presented instead of static results, the assumption can be changed and the entire study instantaneously updated to reflect the new parameter.

The utility of parametric methods for improving the ability of designers to update studies has been demonstrated in aircraft design and in systems-of-systems design reviews. However, these methods have not penetrated the defense acquisition system to a large degree, where static milestones still dominate the process. The potential of these parametric methods to replace the static milestones is immense, and would result in a dynamic product that could be updated rapidly as information about the maturation of technology, the shift in enemy set, or the conditions under which the alternatives would operate becomes more concrete.

3.5.6 Compare Results

The third area of advancement, regret analysis, is also cross-fertilization from long-term policy analysis and provides a way to compare the alternatives that are being considered across the many possible futures.

3.5.6.1 Regret Analysis

Regret analysis is a way of measuring the merit of a particular system solution for a set of operating conditions. Kayne defines regret analysis as “the difference between some choice and the best choice for a particular realization of the uncertainties.” [130] Regret is a fairly intuitive concept for engineering that translates well to the generally accepted definition of regret. Webster’s dictionary specifically refers to a feeling associated with a loss or error. [4] If a regret analysis were conducted for a current situation, the regret would correspond to the difference between the system on hand and a system optimized for the current situation. If the analysis is conducted at the beginning of a particular program, its purpose is to look at the way a candidate solution performs with respect to other possible solutions for a certain future operating condition.

The way the difference between the candidate system and the system optimized for the particular set of future operating conditions is quantified depends on the problem at hand. In most system-of-systems problems, many metrics of interest exist for deciding among candidates, including various measures of performance and cost. It is important, therefore, that the method of measuring the difference between solutions includes all of the measures of merit and weights them appropriately.

A common form of regret analysis is the minimax approach. In essence, minimax strives to find an optimum that is defined by the solution that displays the smallest maximum regret over the future conditions considered. [251] By minimizing the maximum regret, the designer is taking a very pessimistic approach and assuming that the worst possible conditions for system performance will occur in the life of the system. The minimax algorithm is also independent of the likelihood of any future condition. Because system designers usually have some understanding of the most likely future operating conditions, they have more information that should be included in the assessment of concept alternatives.

If information about the likelihood of the various future operating conditions was included in the regret analysis, a more complete understanding of the merits of particular system alternatives. Especially in situations were certain operational conditions are "must haves" and others are "wants," including additional information future operational conditions is desirable. Using techniques such as MCS coupled with regret analysis to explore the system behavior in a variety of future conditions has the potential to provide more robust solutions by fully exploring regions of likely and less-likely operating conditions, and factoring that likelihood into the decision. An additional suggestion for improving the way regret analysis is conducted is presented by Aseeri [15], who suggests normalizing the regret for each candidate scenario. This allows a consistent comparison among systems which may exhibit performance at different magnitudes for different scenarios.

3.5.6.2 Regret Analysis Shortcomings

Regret analysis provides a way to compare alternatives that have some sort of overall evaluation criterion associated with them. However, in the current implementations of regret analysis, the regret associated with an alternative exists only at the discrete points where a scenario has been evaluated. These clusters of individual data points can give the decision maker some sense of the value of each possible alternative, but the discrete nature creates shortcomings in decision making.

Because some futures are generally considered more likely than others, a minimax approach is deficient for decision making because it cannot incorporate the likelihood of the different futures. A possible solution to this would be to assign a weight to the actual value of regret based on the perceived likelihood of the scenario. Unfortunately, this solution only partially addresses the problem. During most decision-making processes, there will be differing opinions on the likelihood of different scenarios, leading to a log jam whenever the regret calculations must be updated to reflect differing opinions.

The minimax approach, even when weighted with the likelihood of the scenario, has the possibility of returning a solution as "best" that is in fact outperformed over the vast majority of the design space. This is illustrated in the comparison of three hypothetical platforms in Figure 16. "Option A" represents one robust design candidate, "Option B" represents another robust design candidate, and "Option C" represents an optimum design candidate focused on performing well in a narrow band in the left half of the scenario space. Using the mini-max approach, "Option B" would be selected as the "best" alternative. However, it is clear that for the majority of the scenario space, "Option A" is a superior alternative, and only is moderately outperformed by "Option B" in a small region of the space. Most decision makers would consider "Option A" superior, but the current regret analysis construct does not allow for this alternative to get fair consideration. If the regret analysis method could be modified to allow rapid consideration of the performance of alternatives in all areas of the scenario space, regret could be considered across all of the space.

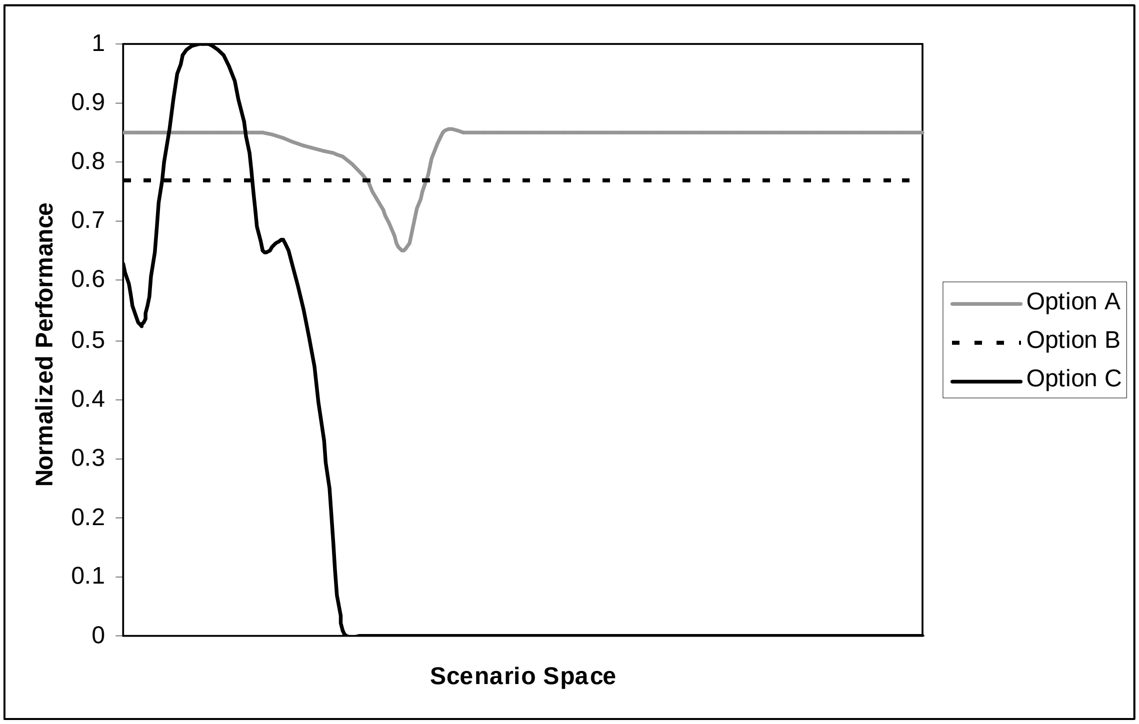


Figure 16: Performance of 3 Platforms

A final shortcoming of regret analysis is that, because of the static nature of the regret analysis approach, it does not naturally fit with the concept of the interactive design review. In an interactive design review the norm is to understand the effects of changing assumptions in real time, which is not possible when a large number of complex analyses must be run.

3.5.6.3 Overcoming Regret Analysis Shortcomings

A more effective approach might be to apply the concept of surrogate models, discussed in detail below, to the scenario space. If a surrogate of the scenario space can be successfully constructed, then because surrogates are merely closed form equations, the regret can be integrated over the entire scenario space, creating a global regret. Global regret is discussed in more depth in following sections.

3.5.7 Tactical Research Question

Can surrogate models be used to address the shortcomings of regret analysis for use in early defense acquisition?

3.5.8 Surrogate Modeling

Surrogate models serve as a way to rapidly assess the results of a particular code, in a particular region of the design space, for conceptual design purposes. In any given area of the design space, the variability of results can be attributed primarily to a handful of variables. While the other variables are necessary for the magnitude of the response, in nearly every case the vast majority of variables can be defaulted within the ranges being considered, significantly reducing the number of computational runs required for design space assessment. This concept, the Pareto Principle, allows designers in early stages of design to concentrate on the design variables that truly matter in the selected concept space.

The identification of significant variables greatly reduces the number of cases that must be considered for a design, but it is often insufficient to allow the real time analysis of trade games and the consideration of multi-attribute decision criteria on the fly. In this case, a surrogate model can be generated based to represent the analysis code in the region of the conceptual design space of interest. These surrogate models reflect the fact that for limited ranges of input variables, analysis codes typically display behavior that can be represented with a polynomial regression equation, an artificial neural network, or a Gaussian Process regression.

Surrogate models are created by careful observation of the analysis code behavior using a Design of Experiments (DOE). DOEs are purposeful manipulation of the significant variables, identified for the particular ranges of interest, with the goal of identifying the effects of each variable and the cross terms between the variables. One of the most straightforward forms of a surrogate model can be generated based on a least-squares regression of the data from the DOE, as outlined in the Response Surface Methodology (RSM) [136], but more complex approaches including artificial neural networks have been applied to various problems.

Surrogate models are simply equations that represent the behavior of a higher-fidelity code or tool with a high degree of accuracy. As continuous equations, they provide the ability to perform more complex mathematical manipulations than pure data from the analysis code would. As equations, they are also platform independent and they cannot "crash" if incorrect inputs are given: the equation itself will always yield a result. The result cannot, however, be relied upon if the input variables are beyond the ranges for which the surrogate was created. Surrogate models have been used to replace a wide range of analysis codes, and can be used for both linear and nonlinear spaces depending on the complexity of the model created. Three common types of surrogates are discussed in the following sections: response surfaces, artificial neural networks, and kriging regressions.

3.5.8.1 Response Surfaces

"Response Surface Methodology (RSM) comprises a group of statistical techniques for empirical model building and model exploitation. By careful design and analysis of experiments, it seeks to relate a response, or output variable to the levels of a number of predictors, or input variables, that affect it." [33]

Response surfaces were introduced in the 1950s by Box and Wilson [34]. The idea behind a response surface is the use of a simple mathematical relationship, such as a polynomial equation (such as Equation 4), to represent a much more complex process.

The coefficients of the polynomial or other equation are often determined through a regression of a set of known data for the complex process. The simple relationship can then be used to find an optimum solution, ideally at a much lower cost than searching the complex process for an optimum.

$$R = \beta_0 + \sum_{i=1}^{n} \beta_i x_i + \sum_{i=1}^{n} \beta_{ii} x_i^2 + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \beta_{ij} x_i x_j + \varepsilon$$

Equation 4

In the 1990s response surface methods began to penetrate into the field of aerospace design for technology assessment through the work of Mavris and Kirby. [159], [136] Integrated Product and Process Development (IPPD), and the Robust Design Simulation (RDS) method both rely on response surfaces to allow statistical design of products and processes. [160]

Response surfaces have some problems when applied to highly complex systems. Because of the typical assumption that the response follows a polynomial equation, discrete responses, nonlinearities, etc can not be captured in the method. This may lead to a sub-optimal solution, since the optimizer uses the (perhaps erroneous) approximation. Additionally, because a least-squares regression is used for estimation of the polynomial coefficients, a non-normal distribution, i.e. fat tails, can lead to a poor model fit. [168]

3.5.8.2 Artificial Neural Networks

An artificial neural network is a type of surrogate model that functions based on the principles of neuron interaction in the brain. [8] Artificial neural networks provide an advantage over response surface equations because they have the ability to capture non-linearities and will work with discrete inputs or outputs. [246]

The basic unit of the artificial neural network is the perceptron, which applies a transfer function to a set of weighted and biased inputs. Most artificial neural networks used in surrogate model applications are composed of three layers of perceptrons: an input layer, a hidden layer, and an output layer. The perceptrons are linked so that the input layer outputs to the hidden layer and the hidden layer outputs to the output layer. [127] The artificial neural network is fit to the data via a stochastic process known as training. Because this training requires a stochastic optimizer, the training must be repeated in most cases to ensure that the "best" possible fit is obtained. [200] The logistics sigmoid function is often used as the equation for a perceptron, and is shown in Equation 5. [225]

$$R_k = e_k \sum_{j=1}^{N_H} \left(f_{jk} \left(\frac{1}{1 + e^{-\left(a_j + \sum_{i=1}^{N} \left(b_{ij} X_i \right) \right)}} \right) \right)$$

Equation 5

In Equation 5, a_j is the intercept term for the j_{th} hidden node, b_{ij} is the coefficient for the i_{th} design variable, X_i is the value of the i_{th} design variable, N is the number of input variables, e_k is the intercept term for the k_{th} response, f_{jk} is the coefficient for the j_{th} hidden node and k_{th} response, and N_H is the number of hidden nodes. [127]

Disadvantages associated with artificial neural networks arise from the computational requirement for training them and from the lack of easily understandable form to the final equation. Because the computational time associated with the training optimized increases with the number of cases, for very large problems training artificial neural networks can be slow. This problem is compounded by the stochastic nature of the optimization, which means that several attempts must be used for each number of hidden nodes attempted in the training process. Also, as the number of hidden nodes increases, the potential for "over fitting" the data increases. Finally, because the artificial neural network relies on the logistics equation, it is often difficult to gain understanding into the physical phenomena behind the system behavior. While polynomial regressions often allow simple linear and quadratic relationships to be identified, the complexity of the logistics equation typically prevents such insight.

3.5.8.3 Kriging Regressions

Another form of surrogate model that has gained significant attention in the past few years is a form of a Gaussian Process called kriging. According to Shao, "a kriging model is a generalized linear regression model that takes the weighted linear combination of a set of collected data as its prediction model." [206] The regression is constructed in such a way as to ensure that the prediction of observed values will always precisely match those values. One of the appealing aspects of kriging is the model's ability to account for a non-linear or multimodal response space. [248], [260] The assumption of a Gaussian Process means that for every predicted point, there is an associated error estimate that is created as a by-product of the model training. [248]

Kriging does not address all issues encountered in the creation of surrogate models. Because of the matrices involved in the creation of the surrogate, kriging methods are not suitable when there are more than 15-20 independent variables or more than 300-500 data points being used to create the model. [132] In these cases the regression becomes too computationally intensive and cumbersome for use as a surrogate.

3.5.8.4 Radial Basis Functions

Radial basis functions are used to build approximations of functions by following the form shown in Equation 6. The radial functions, a function whose value depends on the distance from the center [177], are summed with different centers and weights to approximate the value of the true function. The weights for radial basis functions can be estimated using a least-squares approach or an optimizer in a fashion similar to training a neural network.

$$y(x) = \sum_{i=1}^{N} w_i \phi\left(\left\|\mathbf{x} - \mathbf{c}_i\right\|\right)$$

Equation 6

Radial basis functions provide an advantage over response surface equations for non-linear and non-monotonic spaces. [248] The disadvantages for radial basis function are similar to those of the artificial neural network, in that the equation produced does not provide easy insight into the underlying mechanics of the process and that it is more complicated to produce than the response surface equation.

3.5.9 Establishing MOPs and MOEs

Sound decision making is dependent on the existence of criteria to measure the relative utility of different aspects of alternatives. [181] According to Sproles, an MOE is a metric that quantifies how well a proposed solution meets a particular need of a problem stakeholder. Sproles goes on to draw the distinction between an MOP and an MOP, which is from the point of view of the engineer of other internal developer. [214] The DoD differentiates the levels at which MoEs and MoPs exist, with MoEs existing at a higher level. According to the Joint Test and Evaluation Handbook [63] an MOE is "a quantifiable value that expresses the effectiveness of the system, system of systems, or process under test." The same document describes an MOP as "a quantifiable value that expresses performance or capability of a system, system of systems, or process under a specified set of conditions at the human-machine task level."

Pinker suggests that acquisition metrics can be grouped into six fairly broad categories: cost, acquisition performance, schedule, commercial practices, weapon system performance, and technology innovation. [183] Pinker goes on to outline a large number of metrics that could be used to judge acquisition programs, and while he does not claim that the list is exhaustive, robustness is conspicuously absent. The absence of this metric, which few would argue is unimportant, could be because of the lack of methods for applying robustness early in the defense acquisition system and the lack of an appropriate mathematical definition for judging the robustness of competing alternatives.

3.5.9.1 Global Regret

Measures of effectiveness currently in use for the evaluation of alternatives do not attempt to quantify the robustness of candidates. At most, they observe the change in a variety of MOEs and MOPs for a handful of off-design scenarios.

A new measure of effectiveness is proposed. Global Regret should be a function of the local regrets, regret in conventional regret analysis, across the entire scenario space of interest and of the probability of a certain scenario. The mathematical formulation of Global Regret will be left for the hypothesis testing section of the dissertation. The goal, however, will be to establish a new MOE that will allow the quantification of the robustness of candidate alternatives relative to the other alternatives in the pool of consideration.

3.6 Methodology Summary – Global Regret Analysis

The following steps provide a high level summary of the steps necessary to complete Global Regret Analysis, which will allow the quantification of the relative robustness of candidate acquisition alternatives. Figure 17 summarizes the high-level tasks of Global Regret Analysis.

1. Establish Global Regret as a primary metric
2. Create parametric scenario
3. Analyze alternatives across a wide range of possible futures
4. Create surrogates of the local regret across the possible futures
5. Establish likelihood for ranges of possible futures
6. Evaluate Global Regret

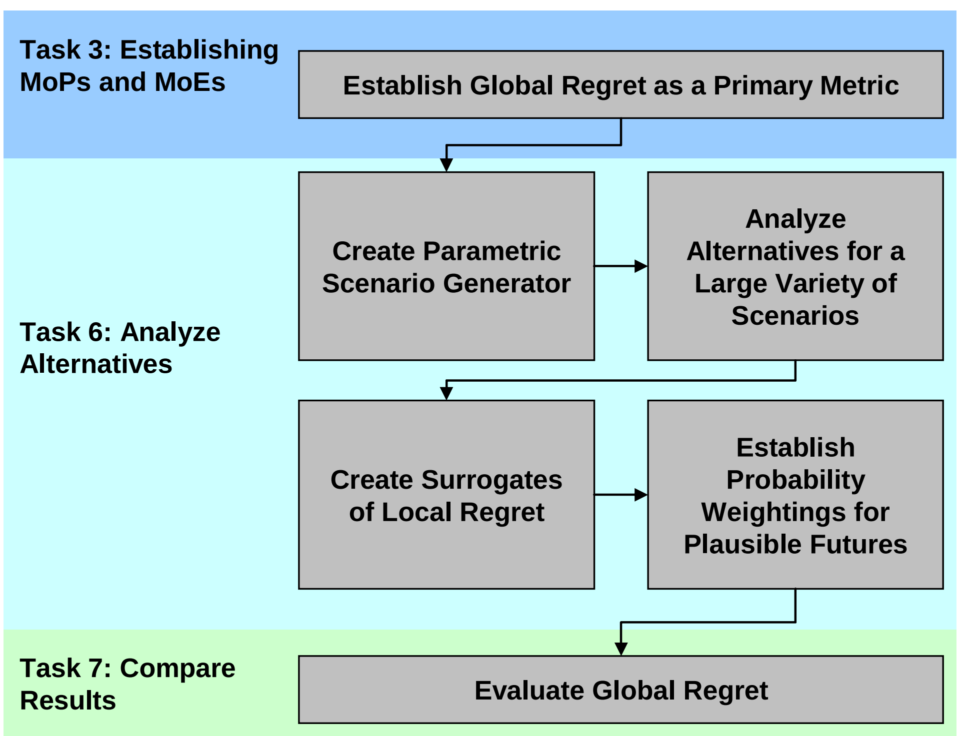


Figure 17: Modification to Regret Analysis

3.6.1 Methodology Information Flow and Tasks

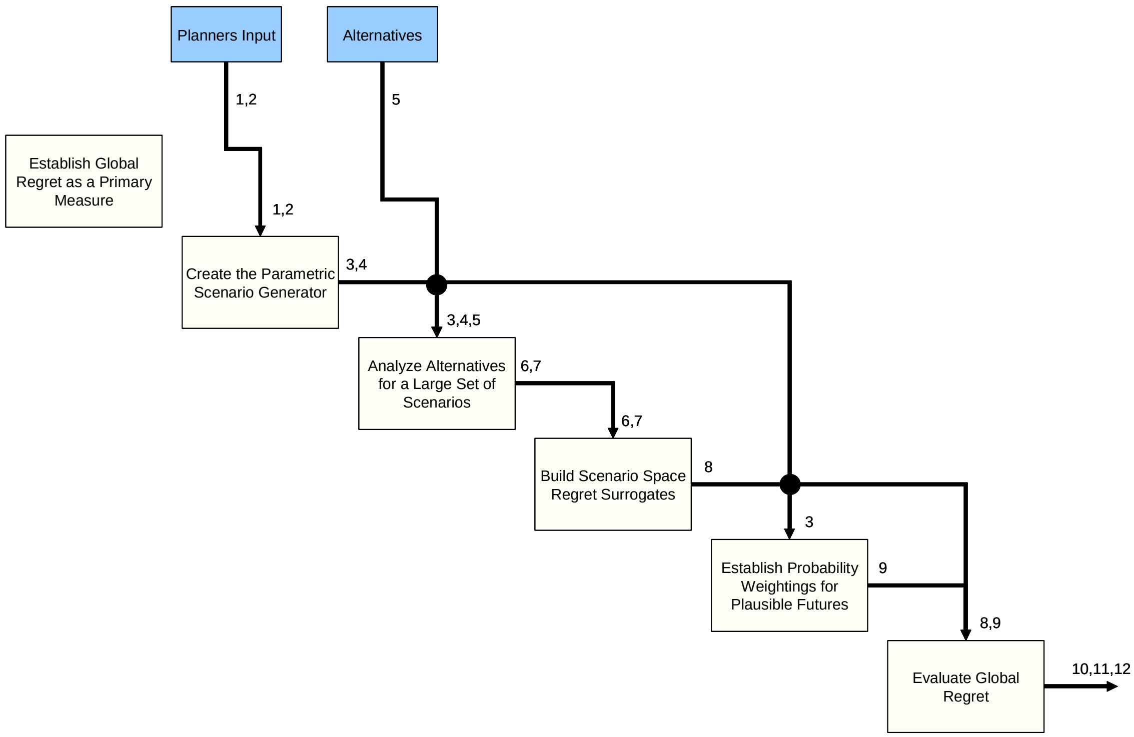


Figure 18: Global Regret Methodology

Figure 18 shows the flow of information among the tasks of Global Regret Analysis. The two blue boxes in the figure show inputs from other tasks in the Pre-Milestone A defense acquisition process, but are not part of Global Regret Analysis. The white boxes in the figure are the 6 activities of the Global Regret Analysis Methodology. The individual pieces of information that flow are coded by numbers and explained below and the tasks for each step are shown in Table 8.

1. Information about the scenario space of interest
2. Information about the expected development of adversary's technology
3. Software tool that can evaluate a concept in a particular scenario
4. Set of variables that can be used to describe a particular scenario
5. Set of alternatives to be considered
6. Scenario/alternatives DoE

7. Results of scenario evaluation for each DoE case
8. Surrogate model defined local regret landscape
9. Likelihood function for scenario space
10. Global Regret value for each alternative
11. Understanding of regret landscape
12. Parametric environment

Table 8: Global Regret Analysis Tasks

Phase	**Tasks**
Establish Global Regret as a Primary Metric	− Ensure robustness is a focus of the engineering efforts
Create the Parametric Scenario Generator	− Breakdown scenario to sufficient level for modeling approach (agent based, discrete event, system dynamics, etc) − Identify interactions significant to metrics of interest − Identify variables necessary to model interactions − Model interactions based on physics, empirical data appropriate for ranges of scenario variables • Should focus on batch mode execution • Should allow for wide ranges of scenario variables

Phase	Tasks
Analyze Alternatives for a Large Set of Scenarios	− Build a scenario/alternatives DoE − Execute the scenario/alternatives DoE using the Parametric Scenario Generator Software − Record metrics for each case in the DoE
Build Scenario Space Regret Surrogates	− Build and evaluate fitness function based on tracked metrics − Fit surrogate models for local regret and other metrics
Establish Probability Weightings for Plausible Futures	− Establish likelihood distributions for each scenario variable − Establish likelihoods for each future scenario
Evaluate Global Regret	− Evaluate the Global Regret Function for each alternative • Can us integration form if computationally possible • Can use numerical approximation of integral • Can use Monte Carlo techniques − Explore scenario space using visual analytics

Information about the scenario space of interest is key for developing the Parametric Scenario Generator. The Parametric Scenario Generator is a Modeling and Simulation

(M&S) environment that has been created in a parametric fashion so that variables critical to the metrics of interest can be rapidly manipulated to create scenarios. The Parametric Scenario Generator also executes the M&S for the settings of the scenario variables. In order to create this environment, however, information about the interactions within a scenario and the expected course of adversary's technological and tactical development must be available. This information comes from military planners, intelligence personnel, warfighters, and engineers who must provide input to ensure the parametric scenario provides a realistic representation of plausible scenarios.

The Creation of the Parametric Scenario Generator step results in an M&S environment that can rapidly evaluate a particular scenario, and a set of variables that describe plausible scenarios. Two additional pieces are necessary for creating a scenario space, however: the alternatives to be analyzed and the scenario Design of Experiments (DoE). The set of alternatives for evaluation using Global Regret Analysis are created in the Alternatives Definition task of the Pre-Milestone A activities. These alternatives must be created using good systems engineering practices, such as those of the RDS methodology. The use of solid systems engineering techniques ensures that the alternatives are representative of regions of the design space where promising solutions exist. The scenario DoE is created by the engineers running the M&S codes and fitting the surrogate models (the following step of the methodology) and must be designed in a manner appropriate for the dynamic nature of the scenario space.

The analysis of the alternatives in the Parametric Scenario Generator results in a DoE that is coupled with metrics of interest for each case. This information is then used to fit surrogate models describing the responses for the entire scenario space for each alternative. These surrogates are then coupled with likelihood functions for the different regions of the scenario space for the evaluation of Global Regret.

The establishment of likelihood for different regions of the scenario space allows the local regret of concepts at different scenarios to be weighted. By weighting the regions of the design space, large regret in unlikely regions will not have as big an impact as large regret in a likely region. However, the likelihoods must be established based on sound intelligence information and engineering understanding about the progression of technology. In general, the lower bound for scenario variables is suggested to be set at the lowest value in the current state-of-the-art, while the upper bound should be some improvement on the cutting edge of the current state-of-the-art. Once these likelihoods are established, the likelihood functions are outputs to the final Global Regret Evaluation.

The analysis of Global Regret establishes a single value that can be used for the comparison of the robustness of individual concepts. However, because of its parametric nature, the landscape of the local regrets can also be rapidly understood by manipulating the values of the scenario variables and seeing the impact on local regret.

3.7 Hypotheses

While the methodology outlined above is the result of significant logical effort, the main assertions remain relatively unreinforced. In an effort to solidify the arguments for using Global Regret Analysis, a number of hypotheses have been developed around the key new developments of the method. These hypotheses will be tested in the following chapter in an effort to understand whether or not the method does in fact address the intended issues.

3.7.1 Hypothesis 1 – Parametric Evaluation of Alternatives

> **Recasting the current Analysis of Alternatives process as a parametric evaluation of alternatives that can be updated throughout the Defense Acquisition Process will increase the robustness of systems-of-systems solutions to a changing future operational environment**

Parametric approaches to systems engineering problems have shown to be effective in increasing robustness by allowing the delay of design decisions until more knowledge about the problem is available. Lack of computational resources limited the usefulness of true system-of-systems parametric studies, but recent advances in aerospace systems design have greatly increased the practicality of parametric studies.

A fundamental assumption of the usefulness of parametric studies for increasing the robustness of systems-of-systems is that knowledge of the near future is better than that of the far future, as discussed by Lempert [140]. Therefore, by delaying decisions about military systems until the fielding date is closer, our knowledge of the conditions will be more accurate. With more accurate information, the system-of-systems can be designed to better complete its task. By continuously updating the assumptions wherever uncertainty exists, the magnitude of the uncertainty in the final product can be reduced.

Implicit in the parametric approach to system-of-systems design is that the system concept must be allowed to change to a certain degree, in this case as information about the future becomes available. In order to allow a changing design, design decisions must be made in as late a stage as possible, using parametric studies for as long a period of time as possible. These synthesized, system-of-systems parametric studies seem overly cumbersome at first glance, especially to anyone who has conducted a complex system design. But recent developments in surrogate modeling now allow a designer to greatly increase the responsiveness of the network-of-systems model to near-instantaneous evaluation. While the creation of these surrogate models does take an up-front investment, both computationally and in man-hours, it affords the designer a great deal of freedom in the assessment of the design space.

3.7.1.1 Hypothesis 1a – Expanded FMCS

> **A hierarchical, surrogate-model based environment can be coupled with a filtered, Monte Carlo decision-making technique to evaluate alternatives in the parametric methodology**

Filtered Monte Carlo decision-making has been developed over the past few years and emerged as an approach to top-down design over a large design space. [78], [24] However, the method has not been demonstrated over a large scenario space. This particular method was selected for the evaluation of alternatives because it represents the state-of-the-art in top-down design space exploration, and allows the consideration of many output and input variables. The natural expandability of the filtered Monte Carlo approach could potentially allow a design space to be expanded to include a scenario space, or a scenario space expanded to include a design space.

3.7.2 Hypothesis 2 – Global Regret Analysis

> **Robustness can be defined as a function of the regret associated with a particular future scenario by using Global Regret Analysis**

In the evaluation of competing design alternatives, it is important to understand the relative robustness of the candidates. Because a suitable method for the early defense acquisition process was not found in the aerospace engineering, methods in other fields similar to defense acquisition were examined. Regret analysis, from long-term policy analysis, was identified as a possible method for quantifying the relative robustness of candidate alternatives.

Regret analysis does have some shortcomings that limit its applicability it its current form, however. The formulation does not return a single measure of the relative robustness, rather a large set of relative regrets. These discrete values can give a general picture of the robustness of concepts, but the designer and decision maker may have difficulties distinguishing among concepts that have different areas of strong and weak performance.

3.7.2.1 Hypothesis 2a – Surrogate Modeling of the Scenario Space

> **Surrogate models can overcome the shortcomings of regret analysis for use in the parametric methodology to enable Global Regret Analysis**

Surrogate models have been used successfully in the past to create a continuous space from a discrete set of data points. Because of their ability to simply represent a complex, but bounded space, they can greatly increase the speed and range over which an analyst may explore a process. This simplification should aid in the ability of a continuous space of regret to be evaluated for a single, general robustness metric.

3.7.3 Hypothesis 3 – Parametric Scenario Generation M&S

> **The use of a parametric scenario generator will allow the consideration of uncertainty across a wide variety of possible futures in the parametric methodology**

One of the key requirements for the current formulation of regret analysis is the ability to quickly evaluate a large number of possible future scenarios, so that each of the alternatives may be evaluated in them. A similar requirement exists for the creation of a surrogate model, in that an environment must be constructed in which a Design of Experiments may be executed. While traditional DoE methods have focused on the parameters under the control of the experimenter, because this method focuses on developing a surrogate of the scenario space, the traditional environment must be refocused.

3.8 Summary

This chapter presented the assertions, gaps, research questions, hypothesis genesis, and hypotheses for the dissertation effort.

Assertion 1 is that Improvements to Defense Acquisition Process could improve MDAP performance and, following, assertion 1.1 is that Pre-Milestone A offers great opportunity for impact. Assertion 2 is that using robustness as a criterion for selecting among alternatives will improve SoS performance.

The first gap is the lack of a quantifiable metric for the robustness of a system. The second gap follows as the inability to account for a massive possible scenario space in assessing robustness. The third gap in the current state-of-the-art and the desired state is the difficulty in updating the Pre-Milestone A activities as additional information becomes available about future operating conditions and technology maturation.

The research questions for the dissertation were derived from the gaps identified in the current defense acquisition evaluation of the robustness of SoS alternatives. They were an attempt to capture the essence of the gap in a manner that could be answered qualitatively or quantitatively through the hypotheses. The research questions for the dissertation follow:

- Most Major Defense Acquisition Programs stretch for more than a decade, so how can we evaluate the robustness of candidate system-of-systems solutions while considering the uncertainty associated with:
    - Technology maturation?
    - Possible warfare doctrine?
    - Possible enemy set?
- How can we define robustness to include these uncertainties?
- How can we promote the ability to update the robustness analysis as higher-fidelity information about the system-of-systems' operating conditions becomes available?
- Can surrogate models be used to address the shortcomings of regret analysis for use in early defense acquisition?

The hypotheses were created by the infusion of techniques from long-term policy analysis and advanced aerospace design. These techniques were Computer-Assisted

Massive Scenario Generation, Regret Analysis, and Parametric Methods. The hypotheses were combined into a methodology presented in the chapter, but are summarized here.

Hypothesis 1 – Parametric Evaluation of Alternatives: Recasting the current Analysis of Alternatives process as a parametric evaluation of alternatives that can be updated throughout the Defense Acquisition Process will increase the robustness of systems-of-systems solutions to a changing future operational environment.

Hypothesis 1a – Expanded FMCS: A hierarchical, surrogate-model based environment can be coupled with a filtered, Monte Carlo decision-making technique to evaluate alternatives in the parametric methodology.

Hypothesis 2 – Global Regret Analysis: Robustness can be defined as a function of the regret associated with a particular future scenario by using Global Regret Analysis.

Hypothesis 2a – Surrogate Modeling of the Scenario Space: Surrogate models can overcome the shortcomings of regret analysis for use in the parametric methodology to enable Global Regret Analysis

Hypothesis 3 – Parametric Scenario Generation M&S: The use of a parametric scenario generator will allow the consideration of uncertainty across a wide variety of possible futures in the parametric methodology.

CHAPTER 4

HYPOTHESIS TESTING

4.1 Introduction

Before pursuing a demonstration of a methodology for robustness assessment early in the defense acquisition process, a series of tests will be performed to indicate the soundness of the hypotheses outlined above. The logic associated with the generation of these hypotheses and an in-depth discussion of their elements will be presented in this section. The hypotheses will be tested with a framework based on the Operation Desert Storm.

4.2 Hypothesis Testing Approach

There are four key elements in the methodology that need to be tested before an example problem of the complete methodology is conducted. Because of the expensive and time consuming nature of defense acquisitions, these elements will be tested in the "laboratory" environment of computer-based simulation. Computer-based simulation offers the opportunity to explore complex relationships at a significantly reduced cost. Unfortunately, validation of computer simulation without real-world data is extremely difficult.

Figure 19 outlines the flow of information among the research experiments and the hypotheses. A bottom-up approach was chosen because the higher-level hypotheses are dependent on the success of the lower-level hypotheses. The bottom-up approach allows the identification of any "show-stoppers" before significant effort is wasted validating dependant hypotheses. Experiment four, which is an experiment in logic, is left for the conclusions section.

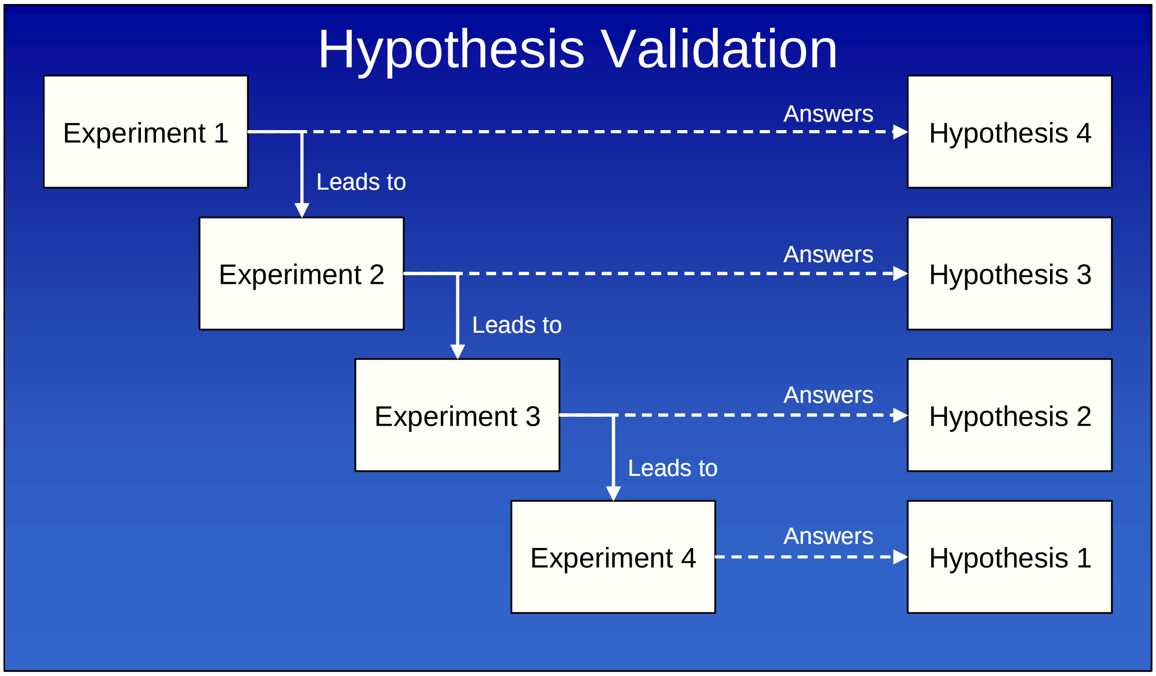


Figure 19: Experimentation Approach

4.3 Experiment 1 – Computer Aided Scenario Generation

Hypothesis: The use of a computer-assisted scenario generator will allow the consideration of uncertainty across a wide variety of possible futures in the parametric methodology

To demonstrate the validity of this hypothesis, a parametric scenario generator will be created in the FLAMES modeling and simulation environment. This scenario generator will then be used to generate a large number of possible engagement scenarios that reflect the physical environment, the threat environment, and the target environment. The scenario generator will be demonstrated by showing a large random generation experiment, a design-of-experiments, and the creation of a particular scenario of interest.

4.3.1 Selection of a Campaign Framework

Before attempting to create a parametric scenario, a set of relevant "ground-rules" need to be established to bound the possible behavior and ensure that the parametric scenario is of use for military analysis. While it might be possible to create a completely generalizable scenario, warfare from the stone-age and the realm of science-fantasy do not particularly lend insight into current defense acquisitions. To test the hypotheses of this dissertation the author chose to limit the parametric scenario to a region around a historical conflict in which strike aircraft played a role. The top row of Figure 20 shows the nine historical US conflicts that were considered, based on the criteria in the first column of the figure. Each conflict was given a qualitative score based on literature search. The criteria are in order of importance, with the most important listed first.

	World War II	Korean War	Vietnam War	Grenada	Bombing of Libya	Panama	Operation Desert Storm	Balkans	Global War on Terror
Data Availability	2	2	2	3	3	2	4	1	1
Technological Similarity	1	1	1	2	2	2	3	3	4
Presence of Air Campaign	4	4	4	2	4	2	4	4	4
Variety of Missions	4	4	4	1	1	2	4	2	4
Historical Proximity	1	1	2	2	3	3	3	3	4

Figure 20: Conflict Selection

4.3.1.1 Data Availability

Data availability was considered the most important aspect of selecting a campaign framework for a number of reasons. The primary reason is the necessity of comparing the results of the modeling and simulation to "real-world" events. Without the ability to make this type of comparison, the experimenter is forced to rely on mathematical proof, or expert corroboration for support. However neither of these provides an attractive way of validation for this dissertation. A secondary consideration is the desire to keep the dissertation unclassified. While a large amount of data exists for any conflict, especially in the second half of the 20th century, much of it remains classified. Publicly available information on military operations remains limited to generalities and statistics, so finding a conflict with as much unclassified information as possible will decrease the likelihood of uncovering sensitive issues.

4.3.1.2 Technological Similarity

Technological similarity is desirable to ease the modeling and simulation workload and to increase the relevance of the results to modern acquisition programs. Current modeling and simulation efforts, for the most part, are focused on current systems of interest. To increase the likelihood of being able to leverage other modeling and simulation work, choosing a campaign framework that is technologically similar is important. Additionally, once the results of the modeling and simulation are analyzed, they will be more likely to be relevant to existing projects if they are based on models of similar technological maturation.

4.3.1.3 Presence of the Air Campaign

The degree of the presence of an air campaign is important because of the focus of this dissertation on aerospace and defense applications. The work of the hypothesis testing has more likelihood of extending to the persistent strike application if it was based on an air campaign from the beginning.

4.3.1.4 Variety of Missions

The variety of missions present in the historical context provides a way to compare many different possible scenarios to historical data. While detailed data about a particular campaign is critical, if that data does not contain a variety of missions the scenario generation can only be validated against a limited number of points. By choosing a campaign with a large variety of mission types, the ability to compare the generated scenarios to historical data increases.

4.3.1.5 Historical Proximity

Historical proximity plays a role in the selection of a campaign framework for a variety of non-quantifiable reasons. Members of the armed forces community have better memory of events that have taken place more recently. Additionally, the closer a conflict is to present day, the more likely that defense planners will take it into consideration when planning from the future. In essence, choosing a campaign framework with close historical proximity increases the likelihood that the work will be applied to current problems.

4.3.2 Campaign Framework - Operation Desert Storm

Based on the scores in Figure 20, Operation Desert storm was chosen as the campaign framework for the creation of the parametric scenario. This choice was relatively insensitive to the weightings of the various criteria for comparison. However, if more data on the current Global War on Terror were available, it would be the campaign framework of choice.

4.3.2.1 Historical Setting

Prior to Operation Desert Storm, Iraq boasted the fourth largest army in the world with nearly a million men. [119] Iraq invaded Kuwait in August of 1990, causing immediate international condemnation of the act. According to the British Broadcasting Corporation (BBC), “On 9 August 1990 the UN Security Council voted 15-0 to declare Iraq’s annexation of Kuwait null and void.” [38] This UN resolution left Iraq essentially out in the cold, occupying Kuwait and waiting to see what the response would come from the West.

After a massive air and sea-lift of military equipment to Saudi Arabia from September of 1990 through January of 1991, Operation Desert Shield, which had been to protect Saudi Arabia from a potential Iraqi invasion, became Operation Desert Storm, to liberate Kuwait.

The coalition consisted of more than thirty nations from around the world, and more than 800,000 troops were deployed at peak strength, with more than 540,000 of those coming from the US at peak strength. [115] Such a large deployment might seem to be overkill in these days of more limited conflicts, but at that time Operation Desert Storm was expected to be a large, peer-on-peer conventional war. Saddam Hussein actually expected to prevail in a war of attrition, as was evidenced by his stubborn refusal to withdraw from Kuwait. [148]

Operation Desert Storm began on January 17, 1991 with a massive effort on the part of the coalition to gain air superiority over Iraq. [101] The initial air-only war lasted until February 24, 1991, when full scale ground operations began. After an extremely short period of ground fighting, the Iraqi army was driven out of Kuwait, and a cease-fire was negotiated on March 1, 1991. The conflict represented one of the shortest in history, but on a tonnage of ordinance dropped per month, the Iraq war rivaled both World War II and Vietnam. [102]

4.3.2.2 Desert Storm Target Environment

The initial phases of Desert Storm consisted of two primary missions: Suppression of Enemy Air Defense (SEAD) and the destruction of known fixed and mobile Scud missile launchers. [Lowry] Throughout the conflict, a wide variety of targets were struck, including fixed and mobile, hardened and soft, dispersed and concentrated.

4.3.2.3 Desert Storm Threat Environment

Three initial attacks on the first night of Operation Desert Storm "created a twenty-mile wide blackened radar corridor for [coalition aircraft] to enter Iraq." [148] This twenty-mile corridor grew in width over the course of the air campaign, and after the first few days of the campaign, the Iraqi air defense threat was limited to both visually and radar-guided Anti-Aircraft Artillery (AAA), shoulder-launched infrared Surface-to-Air Missiles (SAM), and occasional engagements from radar-guided mid to high-altitude SAMs. [148]

Low altitude missions were particularly vulnerable to the radar-guided AAA and infrared SAM threat. In fact, according to the Government Accounting Office, after the first few days of the conflict operations under 10,000 feet were heavily limited during the bombing campaign.

4.3.2.4 Desert Storm Physical Environment

Iraq is a desert country, mountainous in only the north, which is roughly twice the size of Idaho. The summer months in Iraq are typically cloudless [45], which in those months allows for mostly unhindered operation of bombing aircraft. However, there is a threat posed by sandstorms and dust storms, which can cause serious issues with aircraft engine performance. [55]

The mountains of Iraq are nearly 12,000 feet in some areas; however, these regions were not heavily targeted during Operation Desert Storm. The rest of the country, especially the areas in which the majority of Operation Desert Storm's air campaign was concentrated, is primarily flat, offering little in terms of hindrances for aircraft or radar.

While Iraq's summers are nearly cloudless, Operation Desert Storm took place in the winter months, which is Iraq's rainy season. Precipitation itself rarely impacts military operations, but cloud cover can impact the ability to deliver munitions and assess the impact of those munitions. During the winter months, cloud ceilings below 25,000 feet occur 25-35 percent of the time in the areas where the majority of the air campaign was conducted. However, the National Climatic Data Center (NCDC) characterizes the mean cloud cover as "scattered over most of Iraq," during the rainy season. [92]

4.3.3 Creating a Parametric Environment

The parametric environment for massive scenario generation was created by trying to balance a reasonable scope of work for a dissertation while ensuring that sufficient detail was captured to verify the hypotheses. There are a very large number of possible aspects of a scenario that could be considered for scenario generation. However, because the goal of this effort is to demonstrate the hypotheses, a subset of aspects was chosen based on existing models in the FLAMES environment that could be used with some modification. The general categories that were considered were threat characteristics, target characteristics, environmental characteristics, enemy tactics, and friendly tactics. One or several representative aspects were chosen from each area to show how they could be incorporated into a parametric scenario generator. A heavier focus was given to threat characteristics and enemy tactics, as these were considered to be the areas under which the most uncertainty would lie.

4.3.3.1 Selection of a Modeling and Simulation Framework

Table 9 shows four categories of simulation types that were considered for the creation of the models for hypothesis testing. Because of the need to model physics of flight for many of the elements of the simulation, a discrete event simulation was eliminated from consideration. Differential equations were considered inappropriate because, while they work well for systems, the SoS interactions are typically highly nonlinear and complex, which would likely drive the computational expense beyond the level afforded for the dissertation. These eliminations left discrete-time and real-time models. Because there was no need to have human-in-the-loop or hardware-in-the-loop, and the discrete-time models are simpler to program and execute, the discrete-time model was selected.

Table 9: Formalisms of Simulation Types [188]

Formalism	Characteristics	Applications	Issues
Discrete Event	Based on a state machine – time is advanced based on the time of the next state change (event)	Process-related systems or event-oriented systems	Complex systems may contain discrete elements that can use this formalism
Differential Equations	The state of the system varies continuously as a function of time	Systems that respond continuously	Very computationally expensive. Doesn't scale well to large systems
Discrete Time (Time-Stepped)	Simulation time is advanced in fixed and even increments. At each increment, the state of the system is evaluated	Systems that depend on time	Can be slow if there are long periods of little simulated activity
Real Time	Simulation time is periodically synchronized to "wall-clock" time	Human-in-the-loop/ hardware-in-the-loop simulations	Time synchronizing can be complicated

The FLAMES software package from Ternion Corporation was selected as the discrete-time modeling environment for this project. Another package SEAS, was also considered, but FLAMES was chosen because of modeling experience by the author from several research projects. The FLAMES package provided several example models with flight physics, sensor physics, communications, and data recording incorporated. These models were used as the baseline for developing the models used for hypothesis verification. Additionally, the FLAMES environment has the capability to handle terrain, weather, atmospheric properties, and sensor masking, which were considered important in the selection of a modeling environment. Descriptions of the various components of the model follow.

4.3.3.2 Target Components

The target model consists of a single model. It has the ability to be fixed or mobile, allowing for moving targets of various speeds to be simulated.

4.3.3.3 Target Physics

One key aspect of a known target is its ability to withstand an attack, also known as its hardness. A hardened target requires more energy, or a concentration of energy in a very small area, to destroy than a soft target.

The DoD and North Atlantic Treaty Organisation (NATO) define a hardened target as one that is designed to withstand the effects of conventional weapons. They specifically cite the use of rock and concrete as a usual means of protection. [129] Many different types of targets existed in Operation Desert Storm, from relatively soft targets such as aircraft hangars and power plants, to extremely hard bunkers. According to a German designer of one of Saddam Hussein's bunkers under his royal palace is "very, very difficult to crack unless you hit it directly with a small atomic bomb." [16]

Because of the wide variety of targets engaged during Operation Desert Storm, a generic, fixed target will be used throughout the simulation. This target would be representative of a fixed facility, such as a power-plant, command-bunker, or bridge. These types of targets are interesting to military planners and were targeted throughout Operation Desert Storm.

4.3.3.4 Threat Components

The threat for the parametric scenario was broken down into three primary models: a radar model, a tracker model, and a SAM model. After the first days of operation, SAM suppression was relatively effective and reduced the threat to aircraft flying at medium and high altitudes. However, visually and radar guided AAA remained a relatively important threat and caused the coalition mission planners to limit operations within reach of these systems. [26] Because of the limitations on operations below 10,000 feet in Operation Desert Storm, small arms fire and AAA, both radar-guided and unguided, were not considered for this parametric scenario.

4.3.3.5 Threat Component Physics

The following sections give an overview of the modeling approaches used in the FLAMES simulation environment for the various components of the threat for the parametric scenario.

4.3.3.5.1 Radar Modeling

The SAM radar modeled for the hypothesis testing parametric scenario is based on the radar range equation [52], [165].

$$P_r = \frac{P_t G_t G_t \lambda^2 \alpha}{(4\pi)^3 R^4}$$

Equation 7

The wavelength, lambda, of the antenna is based on the operating frequency of the radar. Alpha is the radar cross section of the target, and R is the range of the target from the radar. The gain of the antenna, G_t , is based on the antenna model described below, and only considers the main-lobe gain. The power transmitted, P_t, is user specified. The gain term appears in Equation 7 twice under the assumption that the transmit and receive antennae are one and the same. [222]

For a successful detection, the returned power must be greater than the minimum detectable signal. The minimum detectable signal is defined as the product of the noise at the receiver and the signal-to-noise ratio threshold, which is user specified. The noise at the receiver is calculated based on a product of the noise factor (N_f, user input), the ambient temperature (T_s), transmitted bandwidth (B), and Bozeman's constant (k) (Equation 8). [222]

$$N_r = k\, T_s\, B\, N_f$$

Equation 8

4.3.3.5.2 Tracker Modeling

The tracker model provides the track data necessary from the radar for the SAM targeting and launch. The user can specify the maximum number of tracks, the transmit frequency, and the purge frequency. The tracker is essentially a data repository that purges un-updated tracks at the purge frequency and broadcasts its information to the other models at the transmit frequency. [222]

4.3.3.5.3 SAM Modeling

The SAM model for the parametric scenario is a three degree-of-freedom (3DOF) model that models the translational degrees of freedom. The missile guidance is governed by an expected collision point: the missile flies to where the guidance system "thinks" the target will be located. The missile model is composed of a guidance model and a motion model.

The missile guidance model controls the flight path of the missile. The missile burns all propellant at maximum thrust in the first phase of flight, and then coasts for the remaining flight time. The missile guidance system calculates the path of the target, and the missiles current path to identify an intercept. If no intercept is determined, the missile is turned to the correct path for an intercept. If the difference in the missile heading and the target is more than eight degrees, the guidance system routes the missile on a shortest path to intercept through a maximum G turn. The missile is commanded to explode at the calculated nearest point to the target. [222]

The missile motion model is derived based on Newton's Second Law, that force is equal to the time rate change of momentum [12]]. Missile forces are divided into axial and normal groups (forces are resolved), and the new position is calculated based on the forces. The axial forces include drag, thrust, and weight, while the normal force is the missile lift. Missile drag is calculated based on the dynamic pressure, reference area, and drag coefficient. [222] The use of a simple 3DOF model is consistent with many aerospace modeling and simulation efforts where the details of the missile flight are not the primary concern of the study. [140], [263], [117] The formulation is also consistent with Moore's derivation of weapon performance, which includes information about the "range, time of flight, maneuverability, and miss distance." [164]

4.3.3.6 Physical Environment Components

The physical environment consists of all elements that are not part of the friendly SoS, Target, or Threat sets. This leaves a large portion of the world available for modeling, but it is important to bound the parameters by identifying which impact the elements of interest in other sets in a way that might in turn impact other interactions.

In the Desert Storm Parametric Scenario, two potentially important elements were identified for modeling in the physical environment: terrain, and weather. Proximity to civilians and weather were both eliminated from consideration, however. The weather patterns in Iraq are largely cloudless [92], and all of the systems under consideration are "all-weather."

4.3.3.7 Physical Environment Physics

The physical environment of in the FLAMES simulation impacts many system models. The properties of the atmosphere, both density and wind, impact missile, aircraft, and bomb flight dynamics. Terrain affects lines-of-sight and minimum altitude. Weather can impact line-of-sight in certain ranges of the electromagnetic spectrum (visual, for example) and can limit the aircraft that can be used for a particular mission.

4.3.3.7.1 Terrain Impact

The impact of terrain is primarily manifested in the detection ability of the SAM radar systems modeled in FLAMES. The radar calculation includes line-of-sight, so that the radar detection can not "see" through mountains. A sample detection plot based on a single radar in a mountainous area is shown in Figure 21. The lighter colors correspond to higher elevations. The covered area is represented by the light green region, and is calculated for a red aircraft flying at 300 meters Above Ground Level (AGL) using the FLAMES sensor coverage tool. The red aircraft is shown in the lower right-hand corner of Figure 21. The radial magenta lines from the blue SAM site show where radar coverage was obscured by a higher elevation region.

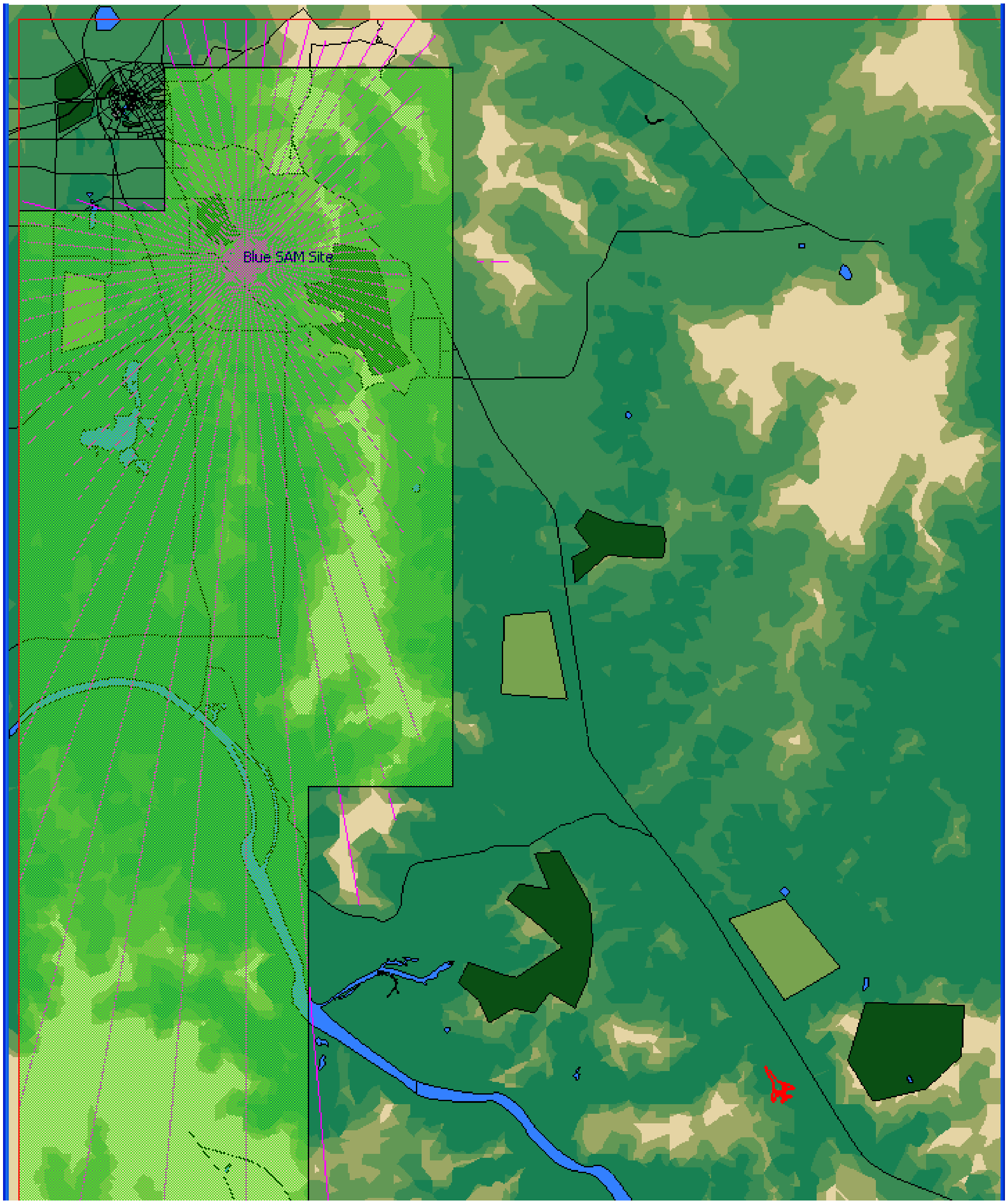


Figure 21: Terrain Impact on Radar Performance

4.3.3.7.2 Terrain Generation

This terrain generation algorithm creates a terrain based on a fractal approach that repeats the basic segment of code over decreasing intervals to generate a “natural” landscape. [69] In two dimensions the progression of the algorithm through 3 iterations would follow the progression in Figure 22. In the top, left-hand area the starting and ending points of a line are specified, in this case at an elevation of 1. In the first iteration of the code, the midpoint of that line is displaced by a random amount elevation (either up or down). This step corresponds to the top, right-hand area of Figure 22. In the second iteration (bottom, left-hand area) the midpoints of each of the line segments created in the first iteration are displaced by a random elevation, but of a reduced magnitude from the first iteration. The final area of the figure shows another iteration of the code. If this were to continue for a number of iterations, the output would look something like Figure 23.

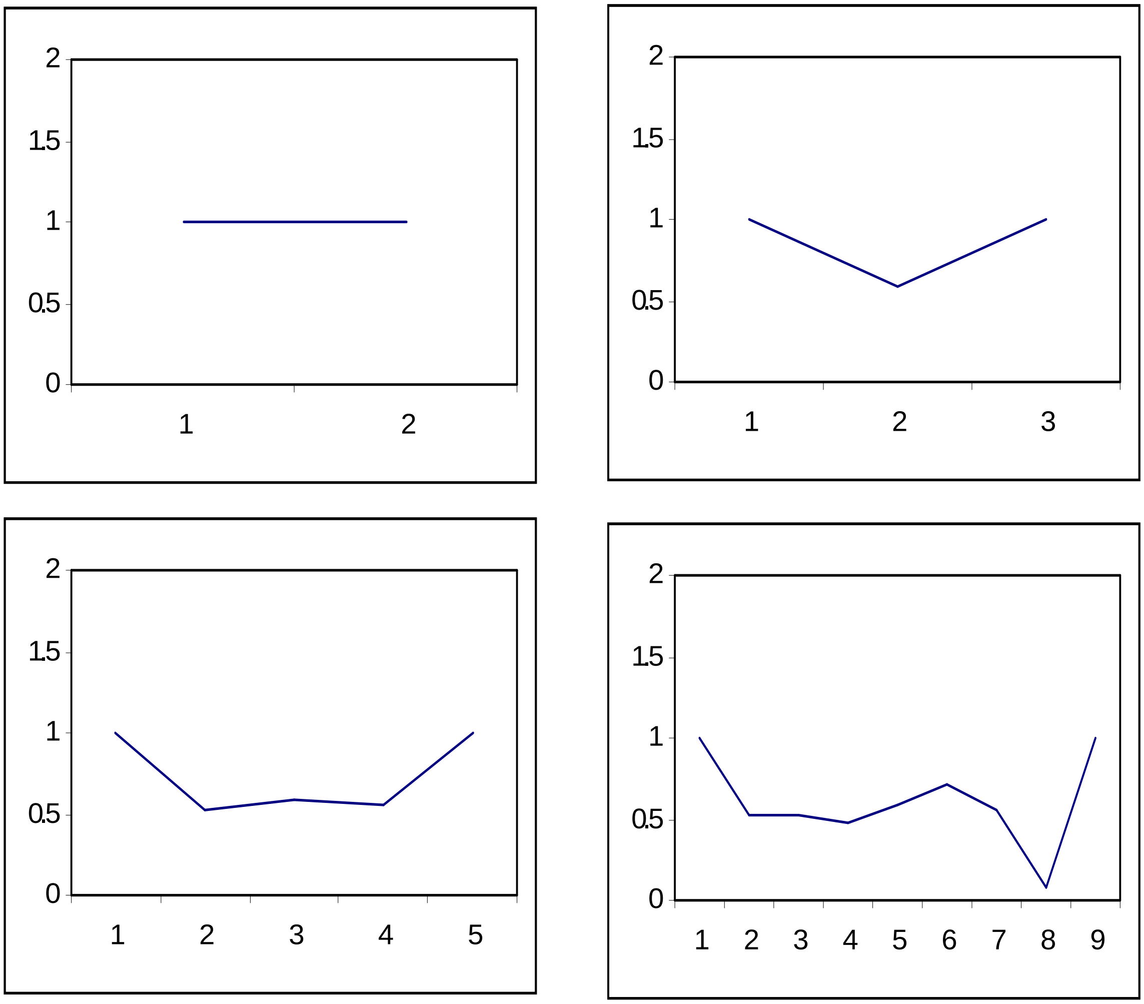


Figure 22: 2-D Progression

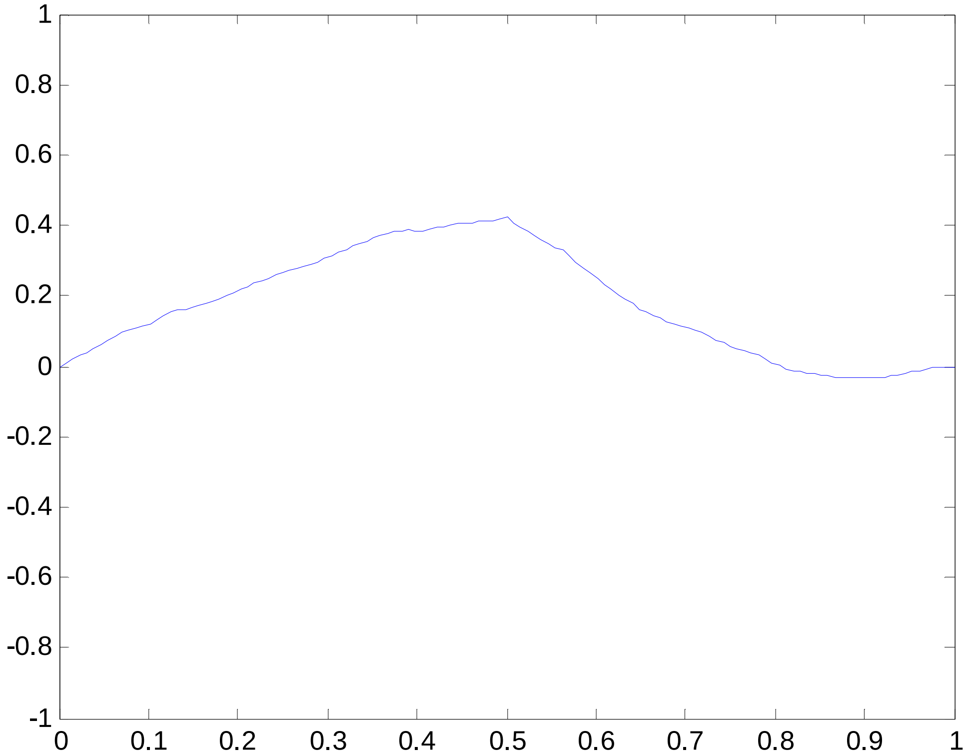


Figure 23: 2-D Ridge

The 3-D terrain generator uses a similar algorithm, extended to rectangles instead of lines. In the start of the code, the four corner elevations are specified. Then in the first iteration, the centroid of that rectangle is displaced by a random elevation, creating four rectangles. Then in the second iteration, the centroids of each of the four new rectangles are displaced by random elevations. This process is repeated until the desired smoothness of terrain is obtained. It is important to note that the computational time for each additional iteration grows exponentially, so using the minimum number of iterations possible is desired. A sample output is included in Figure 24.

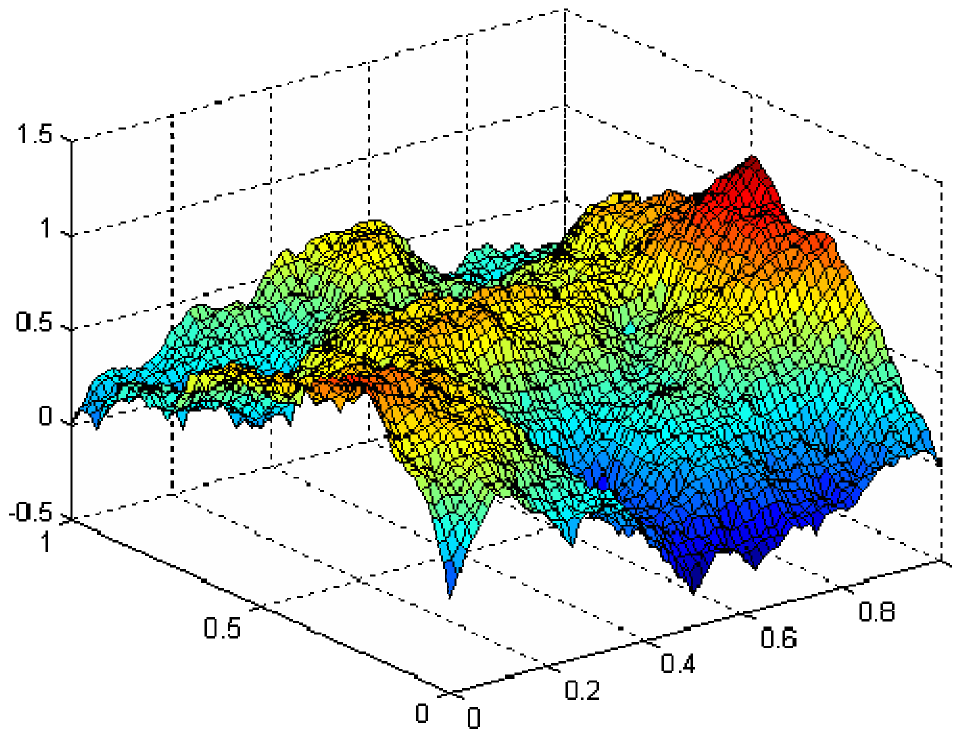


Figure 24: 3-D Terrain

There are three inputs under the control of the user for the terrain generation code. The first is referred to as the "mountainousness" factor. This number allows the specification of the initial range by which an elevation may be displaced. However, if the area is to be re-scaled in a post-processor, this parameter does not impact the results. The second parameter that may be specified is the "jaggedness" of the terrain. This factor allows the user to specify how rapidly the change in elevation decays with each iteration. A lower "jaggedness" factor results in a smoother terrain. Finally, the user can specify the number of iterations through which the code will execute. This increases the fineness of the final mesh.

4.3.3.7.3 Weather

The detailed modeling of weather was considered beyond the scope of this dissertation, and would have vastly increased the computational time required for the execution of the scenario generator. However, the implementation of a cloud density algorithm is very similar to the generation of a random terrain, where areas of high elevation would correspond to areas of limited visibility, and areas of low elevation would correspond to un-hindered visibility. The demonstration of this algorithm was considered unnecessary to the validation of the hypothesis under consideration in Experiment 1.

4.3.4 Scenario Generation

4.3.4.1 Model Center Environment

Traditionally, engineering software has been dominated by command line execution, text based input files, and very limited graphical user interfaces. While usually computationally efficient, these types of legacy software interactions are not conducive to the visual decision making environments that are becoming prevalent in conceptual engineering design. [156] FLAMES can be run from either a command line using an input file for scenario variables, or from the FLAMES graphical user interaction environment, FORGE. [223] In this case, the desire to remain as visual as possible for decision maker interaction must be balanced with the need for rapid execution when exploring massive number of possible future scenarios.

Phoenix Integration's Model Center software provides a good approach to balancing the two competing needs. Model Center allowed the inclusion of both the FLAMES scenario, and the MATLAB terrain generator in a single, visual environment. The environment allows either a single execution of the scenario, or the execution of many scenario runs through the Model Center DOE Tool. [180]

The implementation of the parametric scenario generator is shown in Figure 25. The center window contains the status of the two main programs that make up the parametric scenario generator: MATLAB and FLAMES. There are several additional batch files associated with the parametric scenario generator that handle path issues, deleting database files that are no longer needed, and moving inputs files to the proper directories. Because these files are not affected by input variables, and do not yield any outputs, they are included in the FLAMES icon and have no status associated with them.

The left side of Figure 25 shows the interactive section of Model Center that allows the manipulation of scenario variables and the observation of the results in the scenario outputs. The top four entries are the four scenario outputs that are being tracked. Below those are the 29 scenario parameters that may be changed, as well as non-scenario variables that are discussed in the next experiment. This area allows any input variable to be highlighted and changed, at which point the values in the output will grey out. Once all desired changes have been made, the program icons are activated and the outputs update with the new values.

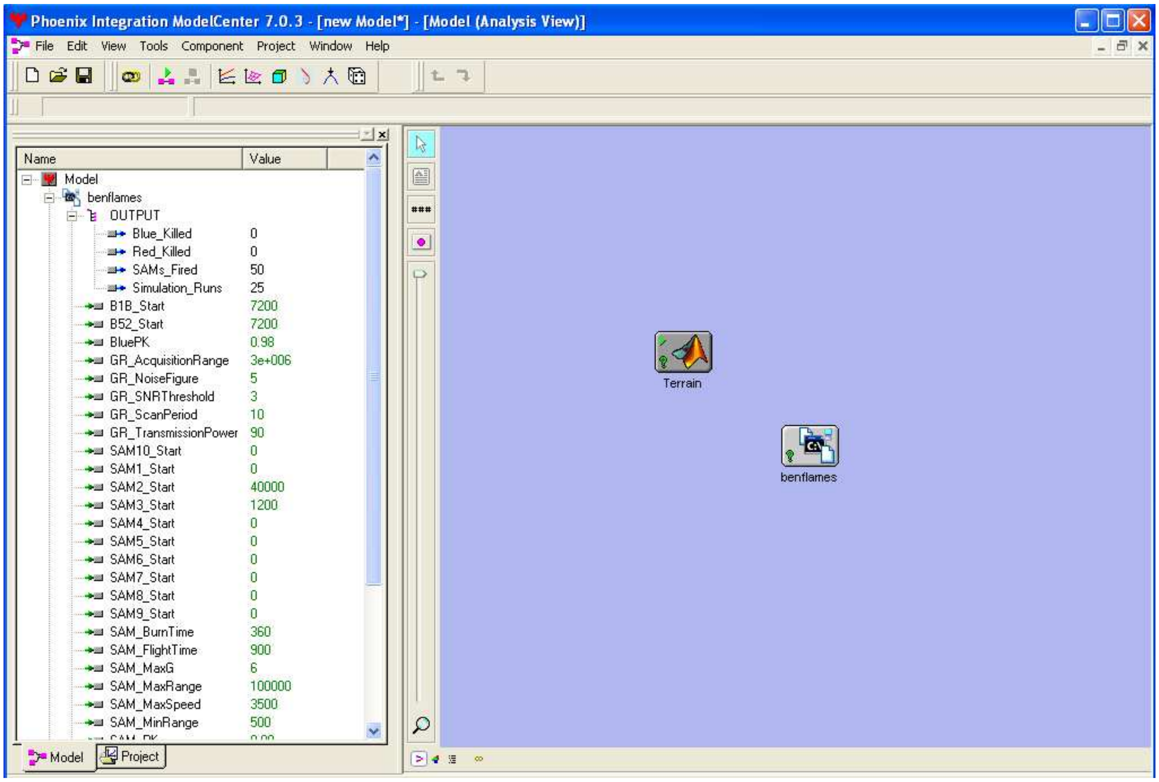


Figure 25: Parametric Scenario Generator

4.3.4.2 Parametric Scenario Generator Function

The parametric scenario generator executes three codes that generate terrain, evaluate the scenario, and parse scenario data. These codes are automated in Model Center so that they perform as a single, stand alone code that returns the values of interest based on input variables.

4.3.4.2.1 Terrain Maker

The function of the terrain maker is described earlier in this chapter. The terrain maker generates an output text file that is read into the FLAMES scenario.

4.3.4.2.2 FIRE

FIRE is the FLAMES executable that actually performs the time-step simulation of the FLAMES scenario. Fire reads in the generic parametric scenario, a variable list generated by Model Center, and the results of the terrain generator. The scenario is then executed 25 times to account for the stochastic nature of many of the variables in the simulation. A variety of key events are tracked and written to an SQL database. These database files include events such as unit kills, munitions firing, unit status, etc.

4.3.4.2.3 FLARE

The database files that are created by FIRE are parsed using SQL script in the FLARE application. FLARE is a command line program that allows manipulation of the data tables and can write results of the manipulations to a text file. In this particular case, blue and red unit deaths are tracked, as well as the number of SAMs fired, and the number of simulation runs (constant at 25). Tracking the simulation runs allows for the easier identification of possible crashed simulation cases. Once the desired data is parsed from the simulation datasets, a batch file erases the data files and transfers the parsed outputs back to Model Center.

4.3.4.3 Ranges of Variables

The table below (Table 10) shows the variables that were considered in the parametric scenario generator. These variables are associated with the terrain generator, the generic air defense system, the ground radar module, the tracking radar module, and the SAM missile module. The first column shows the variable name, the second gives a brief description of the variable and how it relates to the physics-based models described earlier in this chapter, the third column shows the upper bound used for this experiment, and the final column shows the lower bound used for this experiment.

The gray portion of the table shows variables that were not used in the creation of the scenario design-of-experiments. These variables were eliminated to reduce the dimensionality of the experiment and clarify the results. The logic used to select which variables were considered will be discussed in the following experiment, as it directly relates to the metrics used in that experiment.

Table 10: Parametric Scenario Generator Variables

Variable	**Description**	**Upper Bound**	**Lower Bound**
SAM_Dist	Maximum distance by which a SAM can be avoided	120 km	0 km
SAM_PK	SAM probability of kill if target is reached	0.3	0.9
SAM_Thrust	Thrust of SAM burn – drives speed and range in a single parameter	500 lb	5000 lb
SAM_Max G	Maximum gravitational loading for SAM turns	6	30
GR_Trans Power	Ground radar transmission power	40 dBW	60 dBW
GR_Scan Period	Time between a repeat of the ground radar scan	6 sec	30 sec
GR_ SNR	Ground radar signal-to-noise ratio threshold	5	20
GR_Acquisition Range	Maximum range at which the ground radar can acquire a target	100,000 m	600,000 m
SAM_Burn Time	Time for which the SAM burns at full thrust for launch		
SAM_Flight Time	Maximum allowable SAM flight time		

Variable	Description	Upper Bound	Lower Bound
SAM_Max Range	Maximum range for SAM		
SAM_Min Range	Minimum SAM range		
SAM_Max Speed	Maximum speed allowable for SAM		
SAM_Ref Area	SAM drag reference area		
TR_Transmit Freq	Tracking radar transmit frequency		
TR_Purge Freq	Tracking radar purge frequency		
GR_Noise Figure	Ground radar noise figure		
Number of SAMs	Number of active SAM sites in the scenario	0	10
Mountainousness	The maximum elevation change of the scenario		
Jaggedness	How quickly elevation can change between gridpoints		

It is important to not that in most cases, the upper and lower bounds are not fixed values and can be determined based on the size of possible future space that the user wants to explore. The use of physics-based models is important if the bounds of the variables are not defined. An accurate physics-based model should perform for any realistic value of the inputs (unless computational considerations take over), while historical regressions and empirical relationships are only valid in the ranges for which they were created. In this case, because the models are physics based, most variables do not require limits.

4.3.5 Experiment Results

The visualization of a parametric scenario is extremely difficult because of the number of degrees-of-freedom that are in the scenario. Humans have difficulty visualizing any number of dimensions greater than three, so in the following sections, examples will be given in both three dimensions for easier visualization, and in many dimensions that reflect the true capability of the parametric scenario.

The most effective way to present many dimensions of a scenario simultaneously is a multivariate scatterplot. In the scatterplot, each scenario is represented by a point in each box. If a particular point is selected, its vector of attributes is then the values that read on each axis of the scatterplot. In the scatterplot, it is important to realize that each box is a plot of the ordinate versus the abscissa, not the abscissa as a function of the ordinate. Each is independent variables, or if they are dependent, only two dimensions of the problem are being shown in the particular box.

4.3.5.1 User Specified Scenario

The first area of interest for the parametric scenario generation tool is the ability to investigate rapidly a specific scenario of interest, without a large investment of programming time. This type of scenario is useful for decision making exercises where the decision makers want to play "what-if" games about the scenario assumptions. By using the parametric scenario, the decision maker should be able to specify the possible future of interest and then have rapid feedback to the impact of that scenario on the metrics for the decision making exercise.

In three dimensions, this type of scenario might look something like Figure 26. The many possible dimensions of the scenario generator have been resolved to three meta-criteria for demonstration. These generic descriptors of the scenario are SAM Capability, Radar Capability, and Target Defense. If a decision maker were interested in a relatively low-threat environment, but where any missiles present were highly capable, the green circle would provide insight into that type of scenario. The purple square, on the other hand, would provide insight into a scenario where the adversary stationed a large number of low-capability assets near a target. The other two user-specified scenarios represent different combinations of the three enemy scenario parameters.

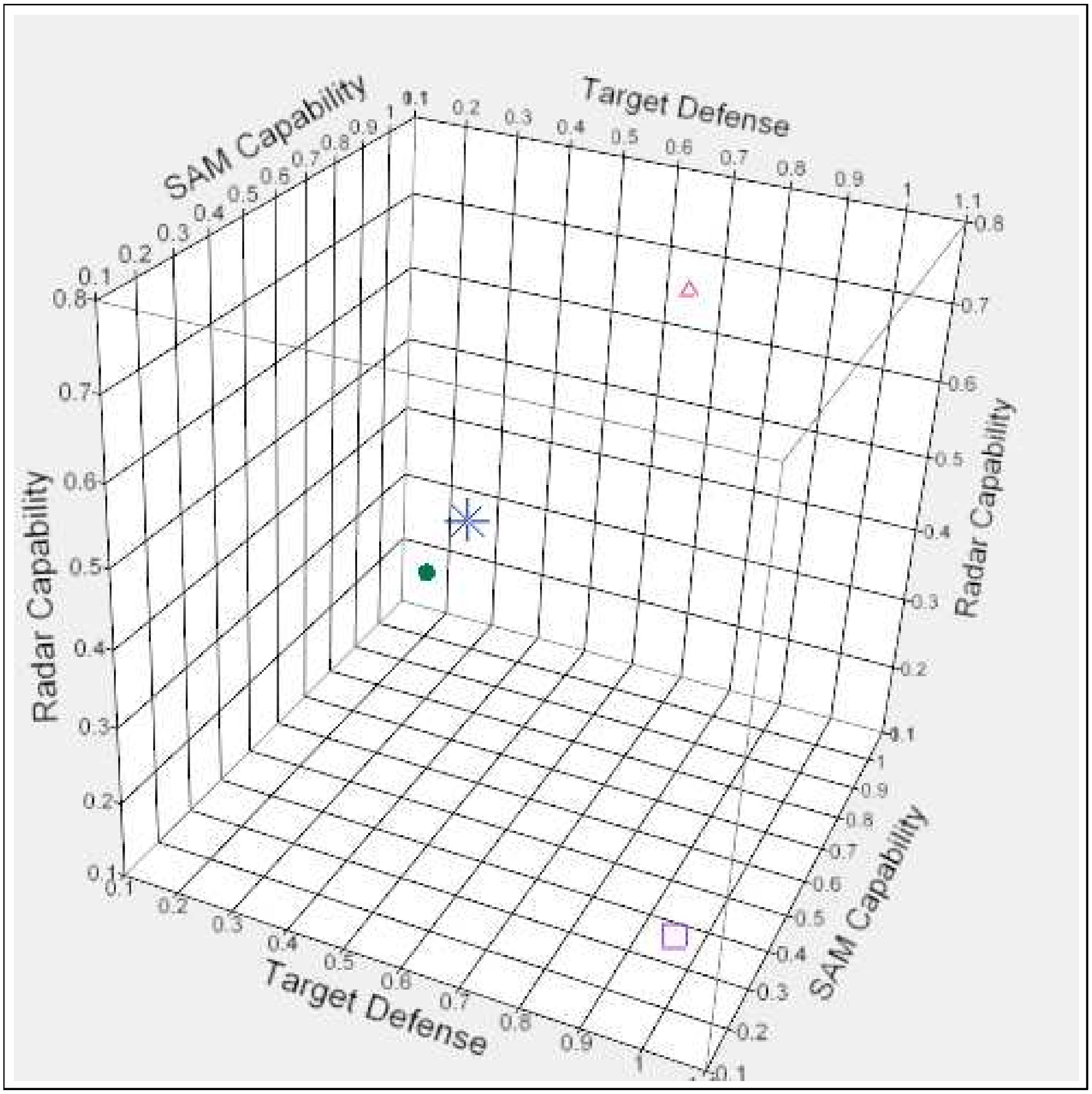


Figure 26: 3-dimensional Scenario Investigation - User Specified

The actual parametric scenario generation tool is much more flexible than the generalized three-dimensional example because of the increased number of parameters that can be manipulated. Table 11 shows the parameter settings and results for a number of different user specified runs, using the parametric scenario generation tool. The blue system being evaluated is a B-1B bomber flying at 10,000 feet and 600 knots. In each case the process of changing scenario variables and evaluating the results took approximately 20 seconds.

Table 11: User Specified Scenario Results

RUN	SAM_DIST	SAM_PK	SAM_THRUST (LB)	SAM_MAX G	GR_TRANS POWER (DBW)	GR_SCAN PERIOD (SEC)	GR_ SNR	GR_ACQUISITION RANGE (KM)	GR_NOISE FIGURE
1	0	0.9	5000	30	60	6	5	3000	5
2	60	0.9	5000	30	60	6	5	3000	5
3	0	0.7	800	20	50	6	5	300	15
4	60	0.7	800	20	50	6	5	300	15
RUN (CONT)	SAM_BURN TIME	SAM_FLIGHT TIME	SAM_MAX RANGE	SAM_MIN RANGE	SAM_MAX SPEED	SAM_REF AREA	BLUE KILLED	RED KILLED	
1	360	900	100000	500	3500	.3143	25	0	
2	360	900	100000	500	3500	.3143	23	4	
3	360	600	10000	500	1000	.3143	19	5	
4	360	600	10000	500	1000	.3143	0	24	

The first run in Table 11 shows a scenario where the blue bomber is forced to fly directly over a SAM site in its ingress to target. This scenario might occur when the enemy has established a heavily defended zone around the target, or when intelligence is lacking for the mission planning effort. Both the SAM and the radar system considered for this scenario are at the upper-end of the capabilities found in the literature search. The radar coverage for this scenario, against a target similar to a large bomber aircraft, is shown in Figure 27. This coverage was calculated by the FLAMES sensor coverage tool. This scenario would be the equivalent of a decision maker asking for a worst-case scenario, and not surprisingly, in each of the 25 iterations, the blue bomber is shot down before it reaches the target.

Figure 27: Case 1 and 2 Radar Coverage

The second run in Table 11 is another attack against a very capable enemy, but in this case, the user has specified that the mission planners will be able to keep the blue bomber at least 60 km from the enemy SAM site. This would correspond to a situation where the intelligence community was able to establish before the mission the likely location of enemy defenses, but some of those defenses were unavoidable. In this scenario, the blue bomber was shot down 92% of the time, but 16% of the time the mission was successful. The difference in these two missions shows that against very capable enemy defenses, mission planning alone is insufficient to create a high probability of success, unless the blue bomber could be kept more than 60 km from the defenses.

The third run in Table 11 shows a scenario where the blue bomber is again forced to fly directly over the enemy air defenses, but with two key differences from the first run. In the third run, the enemy SAM capability is reduced to that more representative of a moderately capable air defense system, the enemy radar system is reduced to a less powerful variety, and the electromagnetic noise perceived by the radar system is greatly increased. The reduction in radar coverage can be seen by comparing the calculated radar coverage against a target similar to a B-52 in Figure 28 to that in Figure 27. In this scenario, the blue bomber survived the mission 24% of the time, and the target was destroyed 20% of the time.

Figure 28: Case 3 and 4 Radar Coverage

The final run in Table 11 shows the same enemy capabilities and noise environment as the third run, but the blue bomber is able to avoid the enemy SAM site by 60 km. In this case, the blue bomber is not shot down in any of the simulation runs, and the target is destroyed 96% of the time. This situation shows a much greater increase in the impact of mission planning on both mission accomplishment and aircraft survival. The difference in the impact of the mission planning suggests that there is a cross-coupling effect on the metrics between the distance by which defenses may be avoided and the capability of the air defenses.

User specified scenarios are not in-and-of-themselves good ways to judge SoS alternatives, especially if only a few scenarios are considered. But, they can provide valuable insight if there is a particular interest in a very limited region of the scenario hyperspace.

4.3.5.2 Random Scenarios

Having established that the parametric scenario generator can be used to rapidly evaluate scenarios of interest to decision makers in a nearly real-time fashion, it is important to see how the generator can be used to sample a large space of possible futures.

The non-dimensional quantities shown in Figure 26 are shown again in Figure 29, but this time are assigned random distributions. For clarity, only 30 points are shown in the space, based on two different types of distributions. The type of distribution assigned to the parameters in the scenario generator depends on the knowledge the designer has about the likelihood of the possible futures associated with the parameter. For example, if the designer has no idea what types of systems the adversary will have, a uniform distribution might be assigned to both SAM and Radar Capabilities, bounded by the range of current state-of-the-art systems and those under development. However, for Target Defense, the designer might assign a triangular distribution, centered on a moderately defended target. This would reflect a perception that the adversary would spread defenses somewhat evenly across possible targets. Figure 29 shows uniform distributions for SAM and Radar Capability, and a triangular distribution centered at 0.5 for the Target Defense.

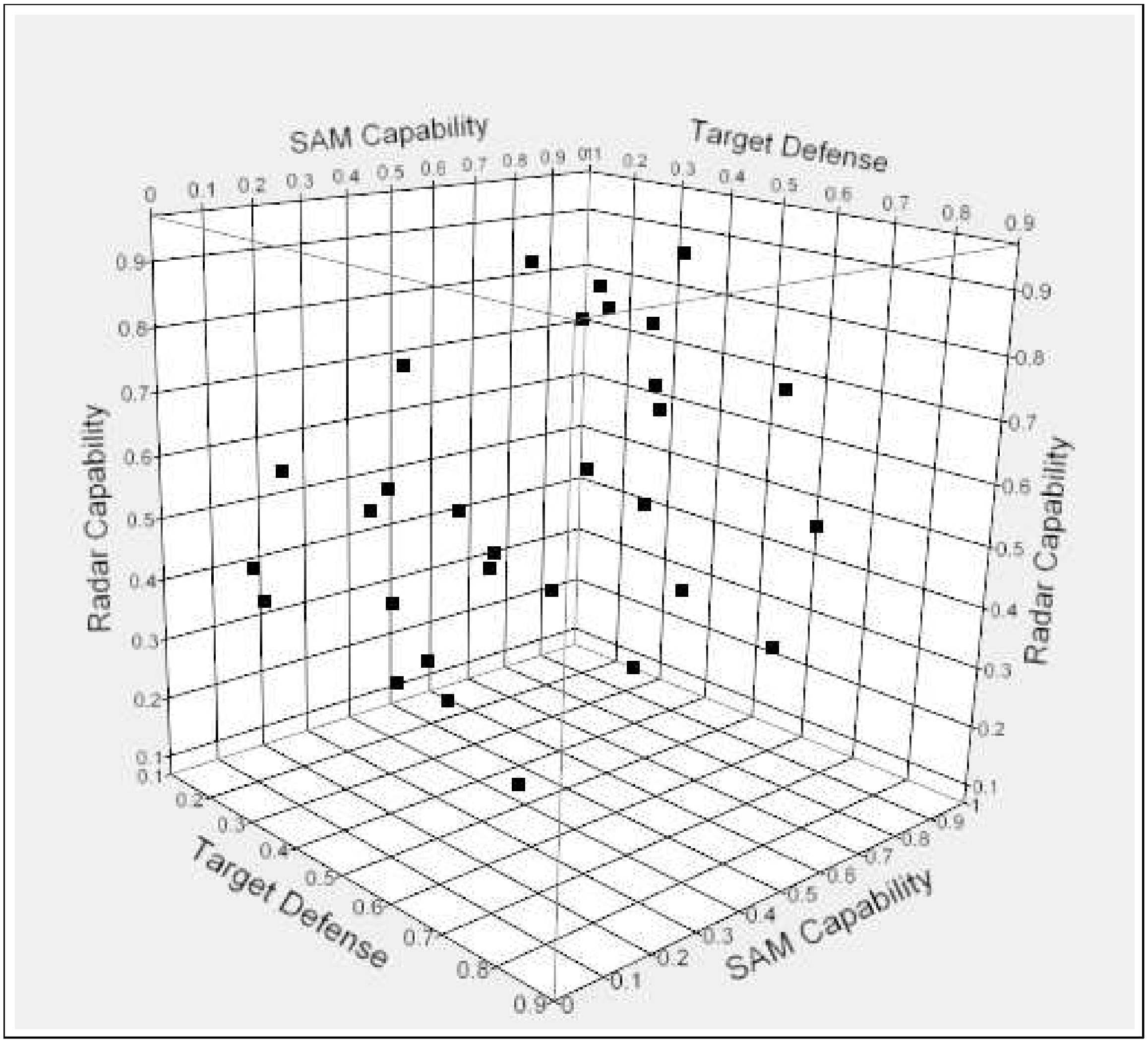


Figure 29: 3-dimensional Scenario Investigation – Random Scenarios

For the demonstration of the ability of the scenario generator to create a large number of possible future scenarios, eight of the scenario parameters were allowed to vary with uniform distributions over the ranges shown in Table 10. These ranges were representative of values found in the literature search on current systems possessed the US and her adversaries. The systems considered were both under development and current state-of-the-art. All information was obtained from publicly available, unclassified documents, so the values may vary from those of actual systems.

Figure 30 and Figure 31 show the range of sensor coverage considered in the random scenario generation experiment. The sensor coverage was calculated based on a large bomber-type aircraft flying at 3,000 meters AGL using the FLAMES sensor coverage tool. [224] The radar-guided SAM systems employed by Iraq during Operation Desert storm fall into the range between the two areas of coverage shown in the figures.

Figure 30: Max Radar Coverage

Figure 31: Min Radar Coverage

The parametric scenario generator was used to create 15,000 random cases. All scenario variables that are not included in the white section of Table 10 were defaulted to representative values. The execution of the 15,000 cases, which represented 375,000 actual scenario executions, took approximately 50 hours of processor time on a desktop computer. The metrics tracked in the scenarios were number of blue bombers shot down, number of red targets destroyed, and number of enemy SAMs fired. Figure 32 shows the results of the random parametric scenario generation in a multivariate plot.

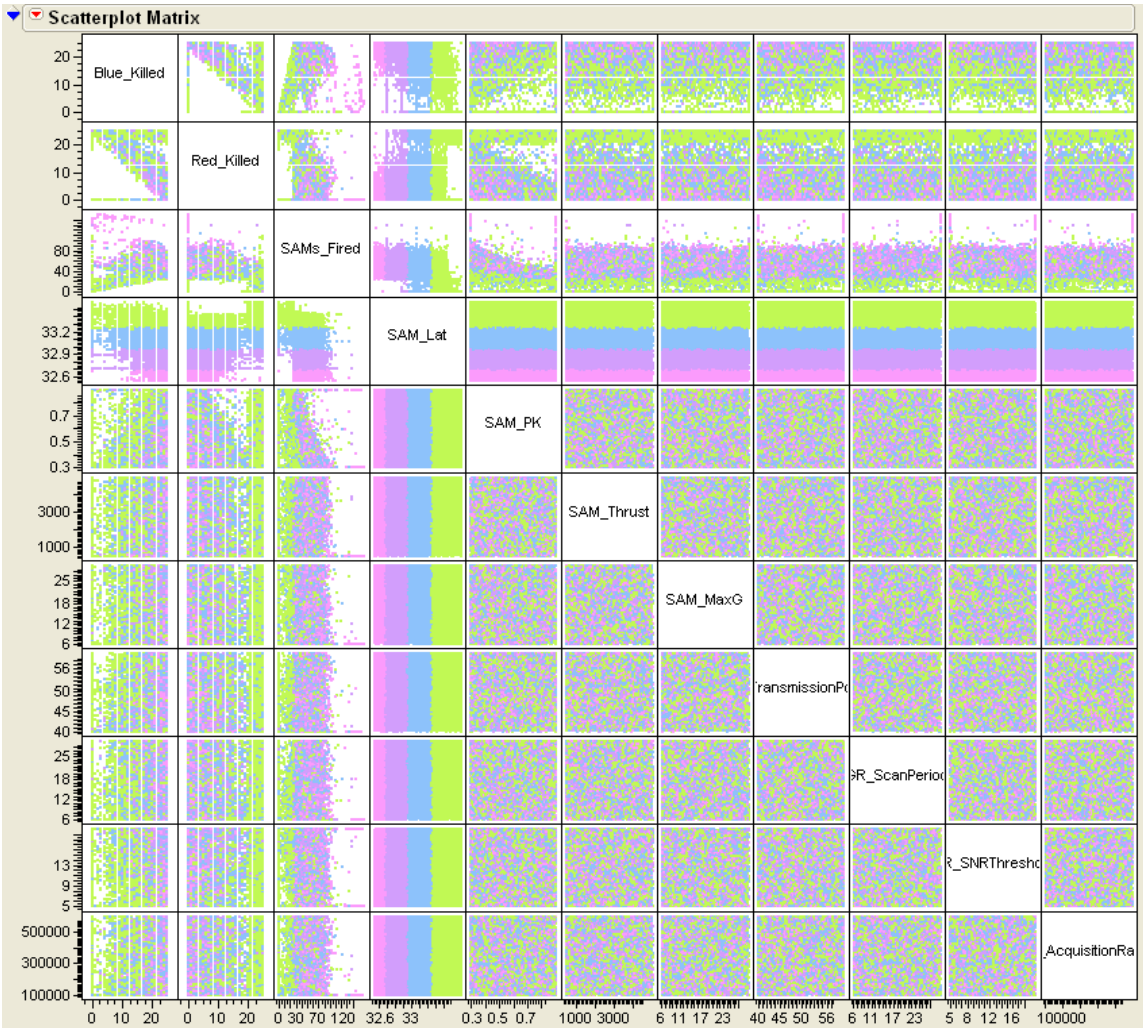


Figure 32: Multivariate Visualization of Random Scenarios

The first three rows and columns of the multivariate scatterplot (Figure 32) show the scenario metrics being tracked, while the remaining eight rows and columns show the scenario parameters that were allowed to vary. A single case will appear in each box of the multivariate scatterplot, and can be traced through by selecting that point in the visualization software. The presence of data points over the total ranges of all the scenario parameters indicates a good sampling of the entire design space. A scarcity of points in a particular region would indicate that a non-uniform distribution was used to populate the scenario hyperspace for that variable. Scarcity of points in regions of the dependent variables (scenario metrics) indicates areas beyond the bounds of the simulation, essentially situations that will be very unlikely to occur given the ranges on the scenario parameters.

To gain insight into the scenario results, four regions of one scenario parameter have been color coded. Green points represent simulations where the blue aircraft was able to maintain approximately 70 km distance from red defenses. Blue points represent a stay-away range of 45-70 km, purple points 20-45 km, and finally pink points are less than 20 km. This color coding does not reveal any dependencies with the scenario parameters in the lower right-section of the scatterplot because no correlations were used among variables. The presence of all 4 colors evenly across the spectrum indicates the uniform distributions. However, interesting trends can be observed in the responses when the different regions of the design hyperspace are color-coded.

Box 1,4 of the multivariate scatterplot (Figure 32) shows the number of blue aircraft killed versus the SAM stay-away range, and is magnified in Figure 33. At this magnification, horizontal lines are present in the scatterplot that reflect the integer values that the loss of blue bombers takes. A trend can be observed in the region of the plot corresponding to a route closer to the air defenses. The lack of points in this region indicates that without mission planning to keep the aircraft away from the defenses, there is a much higher likelihood of losing the blue bomber, regardless of the capability of the SAM site. This relatively intuitive conclusion shows that the scenario generator is producing (at least some) useful information.

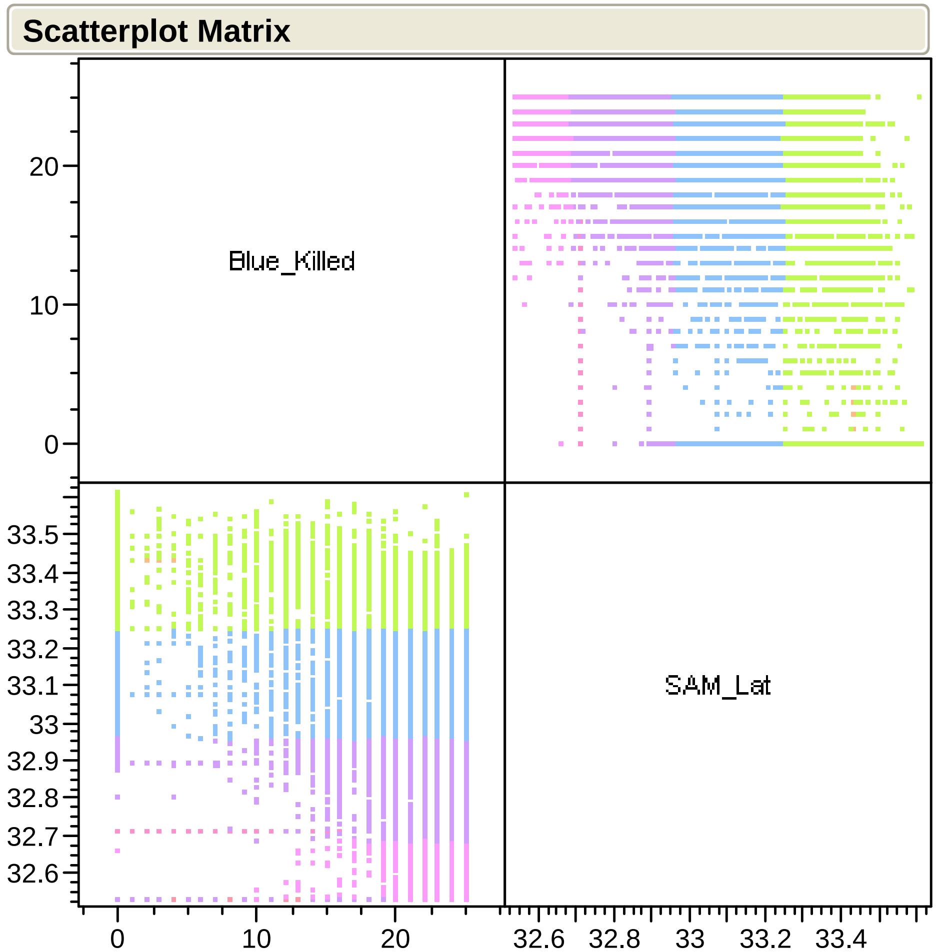


Figure 33: Multivariate Magnification 1

Box 1,5 of the multivariate scatterplot (Figure 32) shows how multiple elements of the scenario can be combined through color coding for additional insight, and is magnified in Figure 34. The change in the presence of colors from the region of high blue losses and low SAM probability of kill to the region of low blue losses and high SAM probability of kill indicates that there is a coupling between the SAM probability of kill and the range by which the blue bomber is avoiding the air defense system. In the top-left (high blue losses and low SAM probability of kill), only the pink and purple bands of range are observed. This indicates that in this region, blue aircraft that were avoiding the defense site were having a higher rate of survival. This observation is consistent with the shift in color down the blue killed axis, showing that low rates of loss occur at further distances. However, as the missile probability of kill increases, the number of points in the region of low blue loss decreases. This indicates that with more capable missiles, the adversary is able to overcome some of the tactical advantage achieved by staying farther from the defense sites. The presence of the green points in the region of high SAM probability of kill and high blue loss validates this supposition.

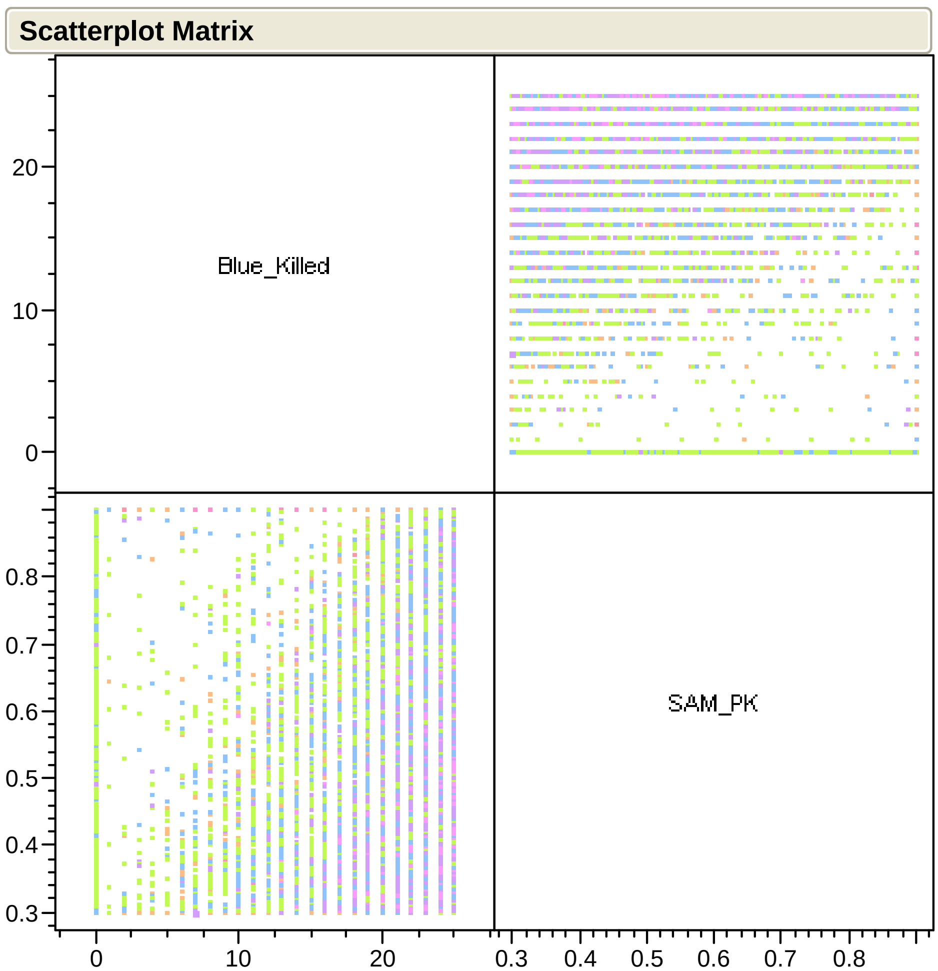


Figure 34: Multivariate Magnification 2

These two examples have demonstrated some of the advantages the multivariate scatterplot has for viewing many possible scenario outcomes, especially when combines with color coding of regions of interest within the scenario parameters. The initial computational investment for creating the data points can be computationally intensive, however, once generated they can be manipulated quickly in the decision making setting. Different views of the data can be created instantaneously, and the color coding and marking scheme can be updated to show different perspectives on the scenarios. The ability to rapidly manipulate an entire scenario hyperspace of data makes the multivariate scatterplot coupled with the randomized parametric scenario generator ideal for conceptual design decision making.

4.3.5.3 Scenario Design-of-Experiments

Just as the parametric scenario generator allowed the creation of a large set of random scenarios, a scenario design-of-experiments can be executed. The design-of-experiments allows a large amount of information to be obtained from the modeling and simulation environment with a minimal amount of computational effort. [See Appendix A]

Figure 35 shows a three dimensional representation of the design of experiments used for the parametric scenario generation. The particular design used for this experiment was a combination design that used a face-centered central composite design, shown by the blue asterisks in Figure 35, and a latin-hypercube space-filling design, the purple diamonds in the figure. The latin-hypercube was selected to have an equal number of points to the number in the face-centered central composite. The face-centered central composite design provides insight into the behavior of the metrics near the edges of the scenario space, while the latin-hypercube points give greater insight into the behavior in the middle of the design space.

As can be seen from the distribution of the points in the three-dimensional plot (Figure 35) the design of experiments provides data points in nearly the entire scenario space. These experiments provide the possibility to rapidly understand regions of the design space where the system under consideration performs well or performs poorly. Once the general performance of the system across the space is understood, greater fidelity can be used to explore regions of interest.

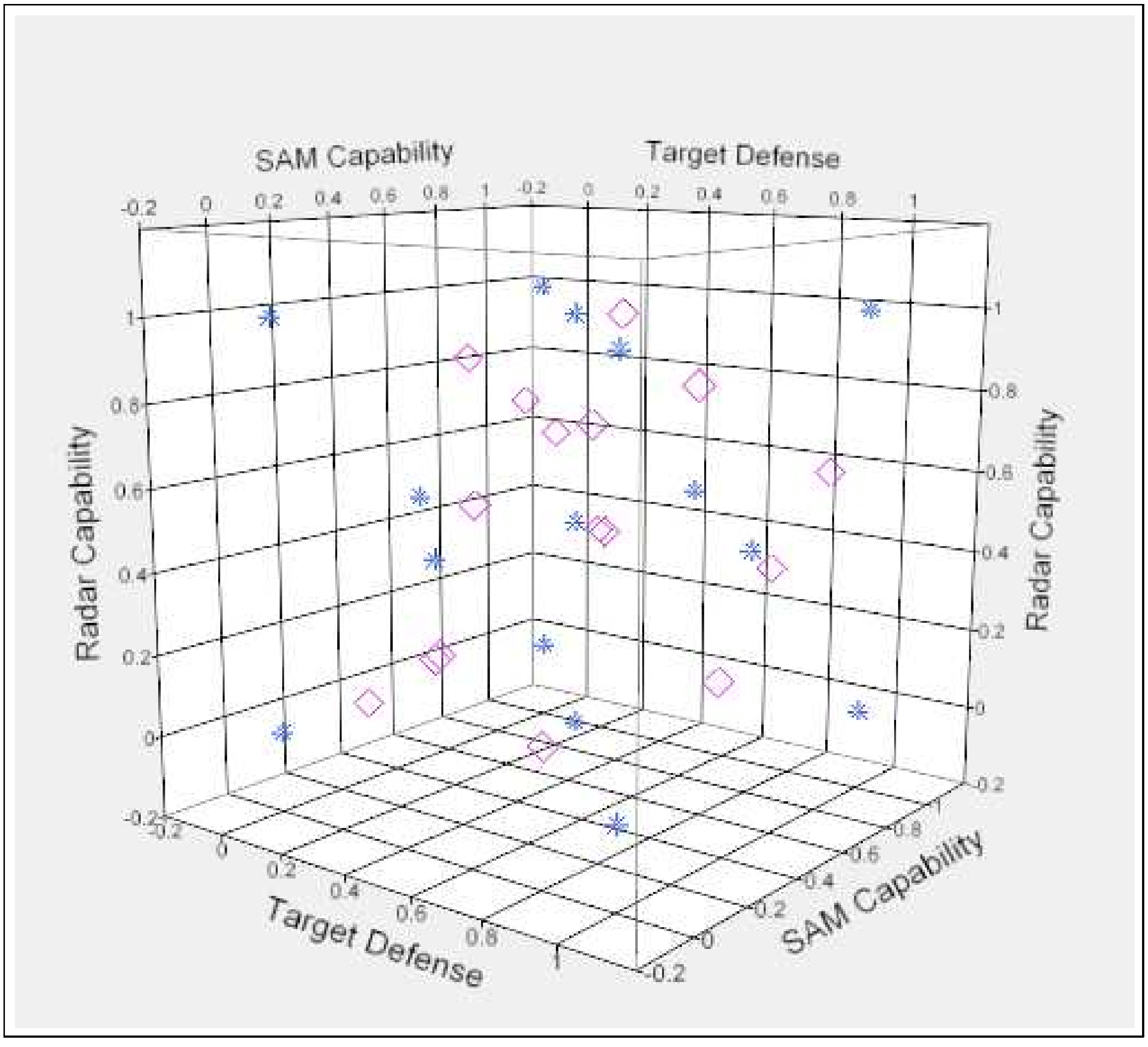


Figure 35: 3-dimensional Scenario Investigation – Design of Experiments

Figure 36 shows the complete design of experiments for the eight variables manipulated in the random scenario generation experiment. In the figure the points have been color coded by the bands of distance that the aircraft maintains from the air defense site. The reason for the appearance of only green points in the bottom right portion of the scatterplot matrix is that at each point there are actually five experiments overlaid. If any particular band were selected, that color would come to the front of the independent variable plots.

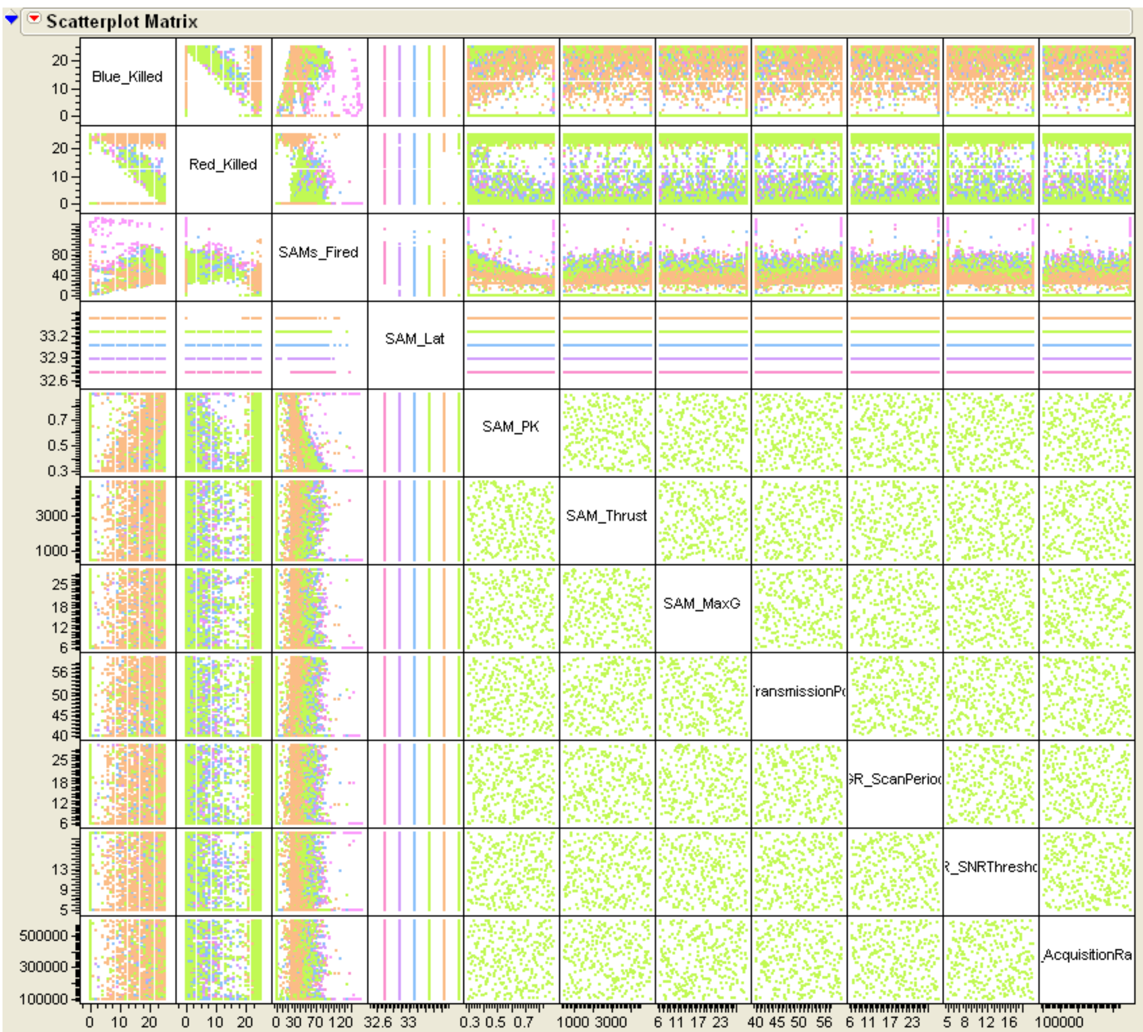


Figure 36: Multivariate Visualization of Scenario DOE

As was observed with the randomly generated points, trends in the location of points and the changes of colors can be used to understand the behavior of the system in the scenario.

Figure 37 shows the design of experiments, but with only one degree of the SAM distance shown (where the aircraft flies directly over the SAM site). The color coding corresponds to four bands of SAM probability of kill. Purple, blue, green, and red points correspond to increasing probability of kill from 30% to 90%.

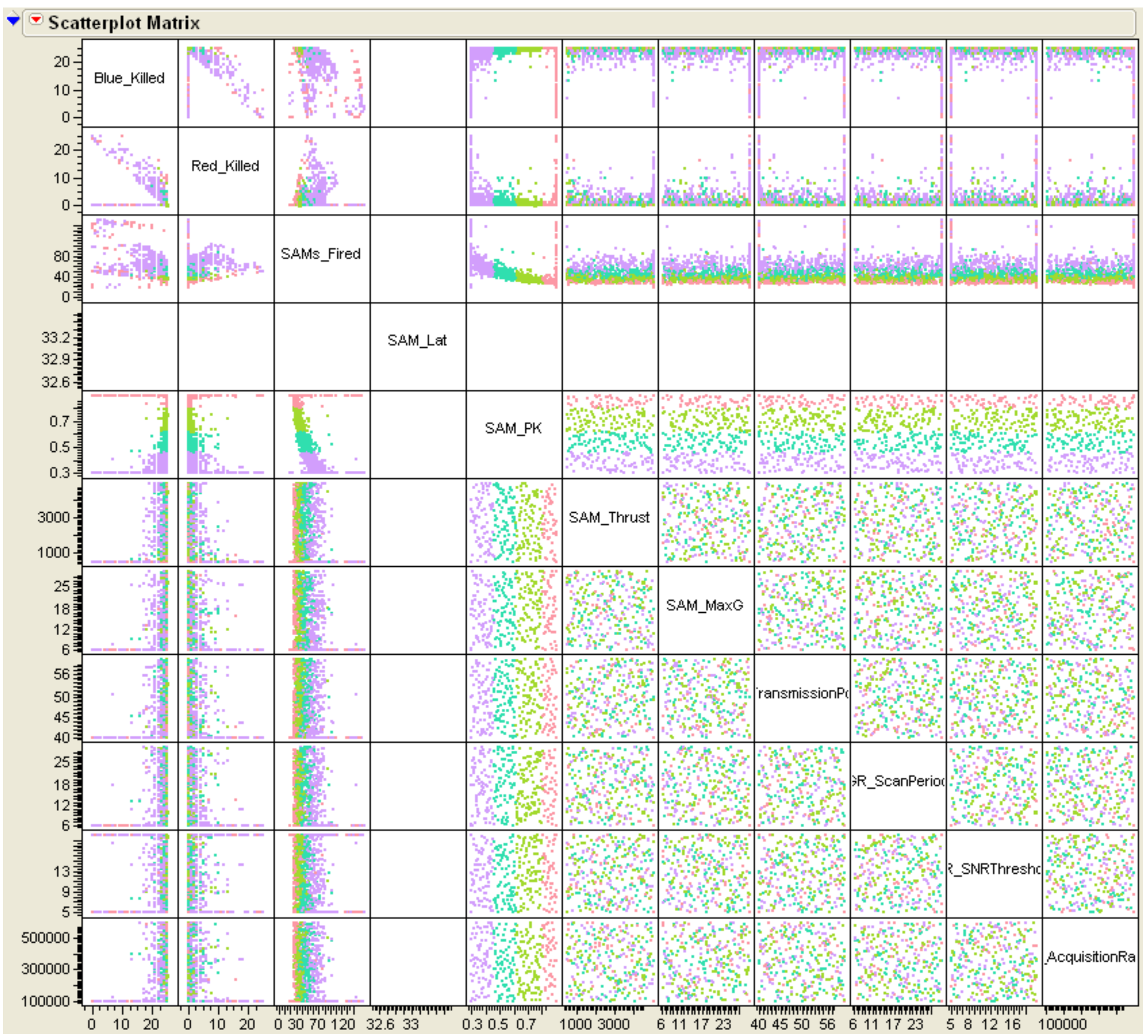


Figure 37: Simplified Multivariate Visualization of Scenario DOE

4.3.6 Experiment Summary

The three components of this experiment explored the use of a parametric scenario generator to evaluate specific regions of interest in the possible futures space, a massive number of randomly created scenarios that span the possible futures space, and finally a design-of-experiments that spans the possible futures space. These scenarios were all used to show trends in the performance of a system can be understood across the futures, even when the uncertainty about the likelihood of the futures was not. In many cases, these trends allowed decisions (such as using mission planning to avoid defenses).

The hypothesis being tested for this experiment is considered true. It has been shown that a wide variety of possible future scenarios can be evaluated rapidly with the parametric scenario generator.

4.4 Experiment 2 – Evaluating Regret

Hypothesis: Robustness can be defined as a function of the regret associated with a particular future scenario.

Sub-hypothesis: Surrogate models can allow the regret for candidate alternatives to be rapidly assessed across the entire future scenario space.

Having shown the possibility to rapidly create a wide range of possible futures using the parametric scenario generator, the next experiment will explore how the robustness of a particular system or system-of-systems can be evaluated across those futures. Four systems that were employed for various missions during Operation Desert Storm will be compared.

4.4.1 Mathematical Definition

The Global Regret, R_G, of a system is the integral, over the possible futures space, of the local regret, R_l, at each possible future multiplied by the likelihood of that future, P_l. (Equation 9) In the equation, the x's are the dimensions of the parametric scenario that may be manipulated.

$$R_G = \int_{x_{1,1}}^{x_{1,2}} \cdots \int_{x_{n,1}}^{x_{n,2}} \left(R_l P_l\right) dx_1 \ldots dx_n$$

Equation 9

The Local Regret, R_l, is the difference in the maximum fitness displayed by a system for a possible future minus the fitness of the system under consideration for that same possible future. The Local Regret is normalized by the maximum fitness displayed by a system for the possible future. (Equation 10) The fitness, F, is a function of the vector of possible future attributes, x_bar, and the system attributes, s_bar. (Equation 11)

$$R_l = \frac{F_{\max} - F_{system}}{F_{\max}}$$

Equation 10

$$F = f(\bar{x}, \bar{s})$$

Equation 11

The probability of a possible future is the product of the likelihood of each element of the vector of possible future attributes. (Equation 12)

$$P_l = \prod_{j=1}^{n} P\left(x_j = \bar{x}(j)\right)$$

Equation 12

This definition of regret is based on the previous definition of regret discussed earlier in this dissertation. The historical definition corresponds to the Local Regret. The development of a Global Regret term grew out of the desire to be able to understand the performance of the system under consideration over the entire possible future space, not just at discrete points. The use of a single, integrated metric for regret allows the use of many decision making methods that would be limited by the existence of many discrete regret data points.

4.4.2 Candidates for Comparison

Three strike aircraft from the USAF's 1991 inventory were selected for comparison in the parametric scenario generator. These aircraft represent a sampling from the large variety of aircraft that were used in the air campaign in Operation Desert Storm. The B-1B bomber is included in this list; however, it was not used during the conflict because of its status as a solely nuclear platform at that point. The desire to show changes in regret as advanced technologies enter a scenario led to the inclusion of this platform, along with the Tomahawk Land-Attack Missile (TLAM).

The aircraft considered were among four used in the heavy bombing campaign. [97] However, because of their similarity in size and capabilities to the F-111, the F-15E was not considered in this experiment. The other aircraft, the F-117 Nighthawk, was only used in a small fraction of the missions and relied heavily on its stealthy characteristics to avoid being engaged by Iraqi air defenses. [97] Because the B-1B represents a "stealthy" platform, and too much information regarding stealth can result in classification of research materials, the F-117 was not evaluated in this experiment. A brief discussion of each aircraft follows.

4.4.2.1 B-1B

The B-1B Lancer (Figure 38) is a multi-mission, supersonic heavy bomber that is the "backbone on America's long-range bomber force." [241] The B-1 program was one of the most controversial defense acquisitions of the second half of the 20th century, but the bomber has become a valuable part of the USAF inventory. [139] During Operation Desert Storm, the B-1B was not used as part of combat operations against Iraq, because at that point it was only armed with nuclear weapons; the B-1B conventional armament program did not begin until 1994. [27] During the weapons conversion program, the B-1B was initially intended for the delivery of Mk-82 non-precision 500 lb gravity bombs, but since has been upgraded to carry a wide variety of weapons including the Joint Standoff Weapon (JSOW) and the Joint Air to Surface Standoff Missile (JASSM). [83] The Lancer was first used in combat in Operation Desert Fox in 1998, delivering conventional munitions. [27]

According to Jane's, the B-1B has a radar cross section of approximately 1% of the B-52. [221] While this value does not give any information about the directionality of the RCS, or the range of frequencies for which it is applicable, the value does give a good starting point for the comparison of the B-52, B-1B and F-111's susceptibility. The use of this value is not meant to generate any real values about the survivability of the B-1B against various radar threats, but rather give insight into general trends between aircraft designed with low-observable considerations and those without.

Figure 38: B-1B Lancer [241]

The parametric aircraft parameter settings used to describe the B-1B are shown in Figure 39. Theses parameters were compiled from publicly available information from the US Air Force's website [241], the Federation of American Scientists [83] and GlobalSecurity.org. [103]

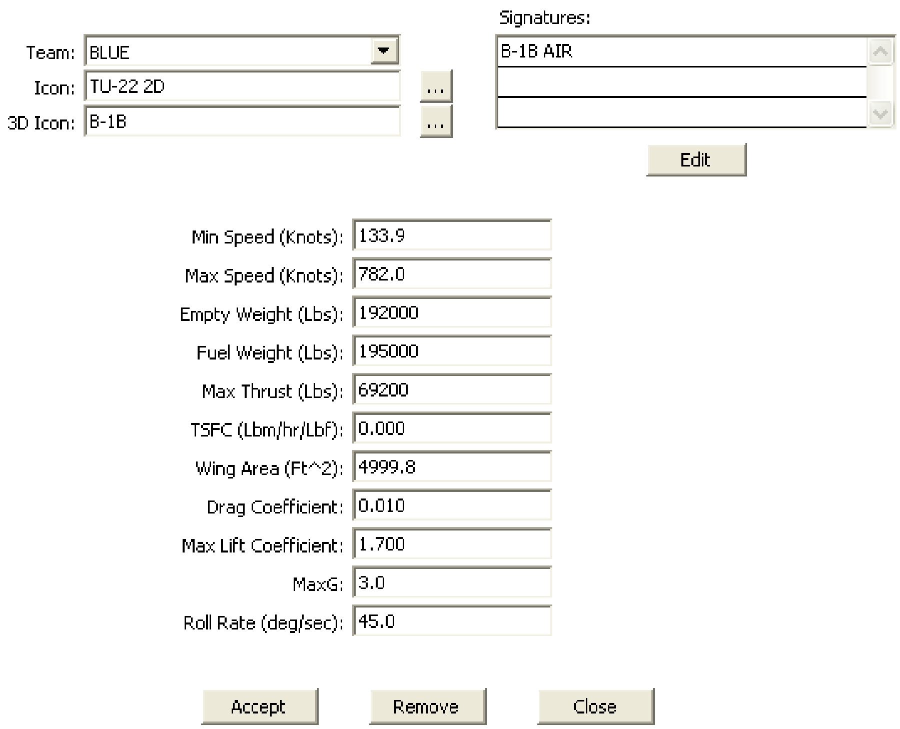


Figure 39: B-1B Parameters

4.4.2.2 B-52

During Operation Desert Storm, forty percent of the weapons dropped by the coalition were delivered using the B-52 (Figure 40). The B-52 was used against a wide variety of targets and is currently able to deploy the most diverse set of weapons of any platform in the USAF inventory. [240] The aircraft is capable of carrying more than 50,000 pounds of ordinance and flying in a range from low-level to around 50,000 feet. [169]

The B-52 has been in the USAF inventory since 1954, and the current generation, the B-52H has been in service since 1961. The B-52 airframe has been continually upgraded throughout its service life, however, and currently has modern avionics systems, global-positioning system, and all-weather capability. [99] The B-52 has a crew of six and a unit cost of $53.4 Million (FY1998). Cost data on an average sortie was unavailable for the B-52 during Operation Desert Storm [97], but will be assumed based on a multiple of the cost of an F-111 mission. Because both aircraft are of roughly the same vintage, the B-52 cost will be assumed to be $100,000 for an average mission. This number incorporates the increased number of crew, and the eight engines of the B-52 (as opposed to 2 on the F-111).

Figure 40: B-52 Stratofortress [240]

Public USAF information and information from GlobalSecurity.org was used to describe the B-52 in the FLAMES model. The parameter settings are shown in Figure 41.

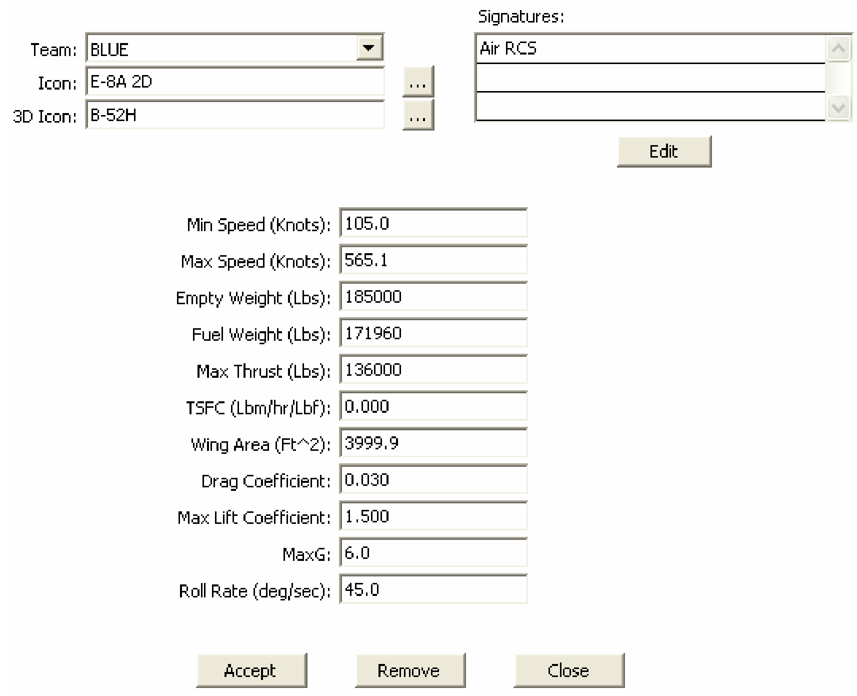


Figure 41: B-52 Parameters

4.4.2.3 F-111

Originally intended as a dual-use platform for the US Air Force and the US Navy, the General Dynamics F-111 (Figure 42) was designed in the 1960's and entered service in 1967. Only the air force variant was built, but the aircraft was designed as a combination fighter-bomber (air force) and air superiority fighter (navy). With a range of nearly 3,000 miles, the F-111 Aardvark filled a long-range, all-weather strike role for the USAF until the last variant was retired in 1998. [170]

The F-111 has a unit cost of $75 Million (FY1998) and is operated by a crew of two. The aircraft saw service in Vietnam, Libya, and Iraq, [120] and was "one of the most effective Allied aircraft in Operation Desert Storm, flying more than 2,400 sorties against Iraqi strategic sites, vehicle formations, and hardened bunkers." [170] The F-111 was initially a controversial purchase, but "achieved one of the safest operational records of any aircraft in USAF history." [81]

The parameters used to describe the F-111 in the FLAMES simulation are shown in Figure 43. These parameters were from publicly available sources including the USAF Museum [170], GlobalSecurity.org [104], and the Federation of American Scientists. [81] According to the General Accounting Office, an average F-111F sortie in Operation Desert Storm cost $24,900 and this number will be used for cost calculations. [97]

Figure 42: F-111 Aardvark [82]

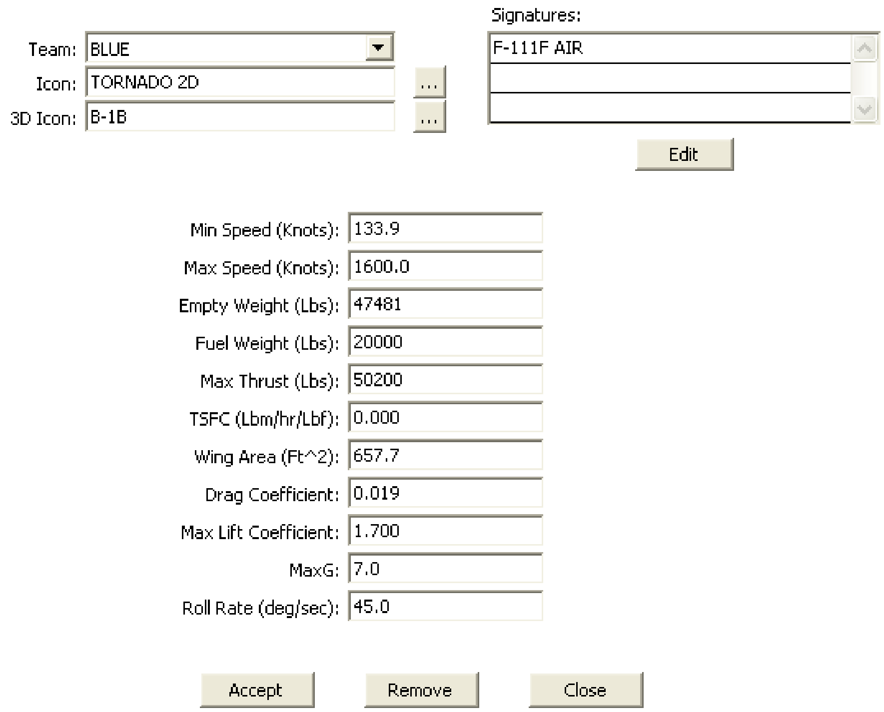


Figure 43: F-111F Parameters

4.4.2.4 TLAM

The Tomahawk Land Attack Missile (TLAM), shown in Figure 44 and also known as the Tomahawk Cruise Missile, is a long-range munition capable of attacking targets 1500 miles from its launch point. The TLAM was initially deployed in 1984 [235], but became famous for its role in Operation Desert Storm as the first shot that was fired in the war. [17] Over the course of the conflict, 288 missiles were launched against a variety of Iraqi targets. [46]

Figure 44: Tomahawk Land Attack Missile [107]

The TLAM's small size and low-altitude flight make it difficult for the adversary to detect the missile in flight. [190] The TLAM also incorporates intelligence about the threat environment and is piloted over an evasive route to its target. [235] The unit cost of the TLAM has varied widely over its deployment since 1984, but according to the General Accounting Office, a TLAM sortie in Operation Desert Storm cost $2.855 Million. [97]

Because of its similarity in flight to a very small aircraft, the TLAM was modeled using the FLAMES aircraft physics model. Figure 45 shows the parameters that were used to model the TLAM. These figures were taken from publicly available sources including US Navy FactFile [235], Raytheon Documentation [190], and GlobalSecurity.org. [107] The flight model completes with the deployment of a bomb munition. The deployment does not destroy the vehicle model, however, so its existence is ignored after the deployment and the deployment itself is counted as the destruction of the vehicle.

After constructing the model, it was decided not to include the physics of the TLAM in the regret consideration, but rather to incorporate it probabilistically. According to the Chief of Naval Operations "about 85% of the 288 missiles fired during the war hit their targets." [46] Because of the dissimilarity between the TLAM and the other aircraft being considered, in size, tactics, and employment, the model used for the TLAM for regret will not follow the pattern of the other aircraft. Instead, the mission success rate of 85% will be used for all of the scenarios in the global regret analysis.

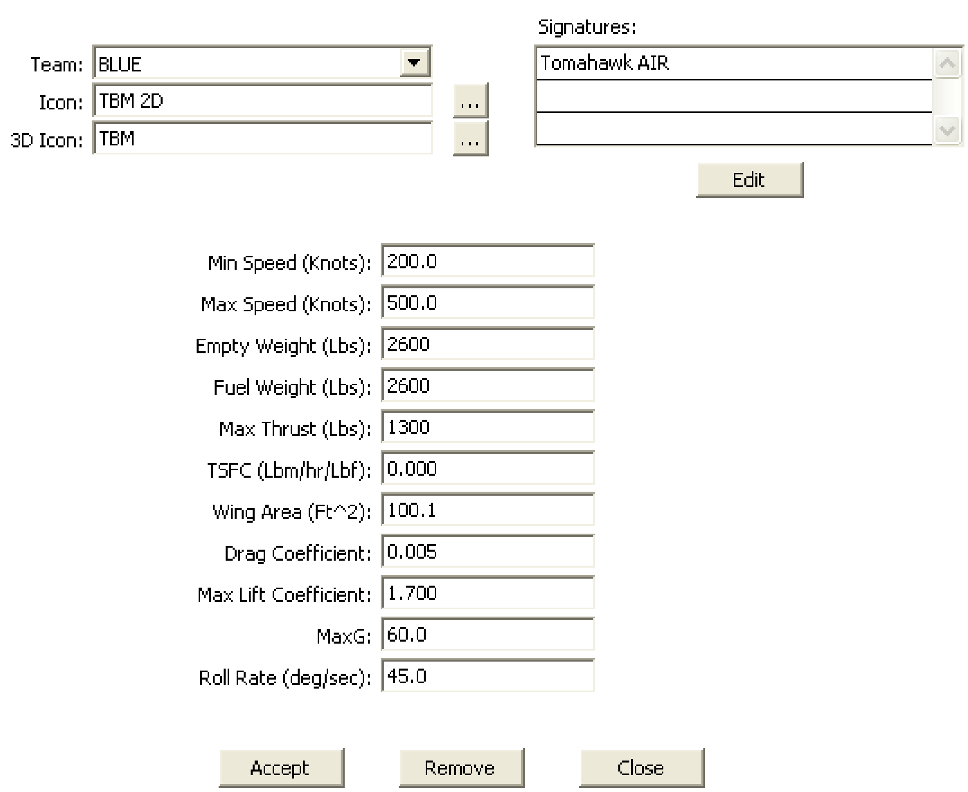


Figure 45: TLAM Parameters

4.4.3 Modeling Candidates

4.4.3.1 Aircraft Physics

The aircraft in the simulation are modeled using a 3DOF translational model. In the parametric scenario generator, the aircraft is assumed to fly a direct path to its target. Because the particular location of the target is not prescribed, different types of flight path are accounted for by moving the location of the target's defenses relative to the flight path. This is equivalent to moving the flight path, but eliminates the need to program a complex flight path for the aircraft.

Biltgen provides a good description of the function of the FLAMES aircraft model used in this effort. [24] The Flames Example Models Documentation [222] also provides additional information. Atmospheric density and dynamic pressure are calculated at each time step for the aircraft. These calculated values are used to calculate the drag on the aircraft, the maximum turn acceleration available to the aircraft (based on lift), the maximum turn rate and minimum turn radius. The maximum roll rate of the aircraft is an input. Within these constraints the aircraft operates based on the movement commands entered into the individual aircraft code.

Because of the mission formulation selected for this experiment, the aircraft flies a straight, level, constant-speed path to the target, then performs a maximum g turn and returns to the starting point of the simulation. The physics of the flight before this ingress and egress are not calculated; they are assumed to not play a role in the scenario level metrics of interest for the regret analysis.

4.4.3.2 Aircraft Flight Conditions

Two primary variables were considered for adjustment to aircraft tactics to ensure that, when comparing the performance of systems, the ‘best’ performance for that particular system could be used.

Aircraft speed was varied from 300 knots to the aircraft’s maximum speed. This range allows the “best” speed for avoiding enemy SAMS to be used. However, the top end of the speed range impacts the accuracy of the bombing simulation. Aircraft flying faster are more likely to miss their targets that those flying at lower speeds. As a result of these two competing attributes, a wide range was considered so that the best trade between survivability and mission success can be made.

Altitude aircraft altitude was allowed to vary from 10,000 feet AGL, to 50,000 feet AGL. This range represents a wide variety of possible altitudes that reflect mission altitudes used in Operation Desert Storm. According to the General Accounting Office, because of the severe threat posed by radar guided AAA and IR SAMS, after day three almost no low-level missions took place. [97]

4.4.3.3 Quantifying Survivability

Electronic warfare (EW) and stealth technologies both play a large roll in the survivability of strike aircraft in hostile airspace. Stealth technologies reduce the RCS of a vehicle, making the vehicle appear “smaller” to radar detection systems. Electronic warfare helps mask the presence of strike aircraft by emitting electromagnetic energy in the frequency range of the hostile radar systems. However, the specifics of both stealth technology and electronic warfare for US systems fall solidly in the classified realm. [175] As a result of the desire to keep this dissertation in the publicly releasable realm, specifics of RCS and EW will not be included in any simulations for this work. However, a general susceptibility term will be established to mimic the effects of some stealth technologies. The following sections discuss the general effects of stealth and EW, as well as how they might be incorporated into a classified study using the methodology outlined in this dissertation.

4.4.3.3.1 Radar Cross Section

Radar systems emit electromagnetic energy and then "listen" for the echo of that radiation off of objects downrange from the transmitter. The amount of radiation that is returned depends on a variety of factors including the propagation effects of the atmosphere, the downrange distance of the object, the material the object is made from, and the size of the object, to name a few. The amount of radiation returned to the receiver antenna is proportional to how easy it is for the radar to "hear" the object.

Increasing the difficulty for enemy radars to detect an aircraft is one approach to reducing the likelihood that an aircraft will be neutralized before it can complete its mission. Survivability is defined as the probability of being detected, times the probability of being shot at if detected, times the probability of being hit if shot at, times the probability of being defeated if hit (Equation 13). A reduction in RCS acts on the P_{detect} term of the survivability equation.

Equation 13

$$P_{Survive} = P_{\det ect} P_{engaged} P_{hit} P_{killed}$$

There are a fairly large number of approaches to reducing the RCS of an aircraft, but the fall into three broad categories: reducing the physical size and shape of the aircraft, coating the aircraft with a radar absorbing paint, and changing the materials from which the aircraft is made to more electromagnetically transparent materials.

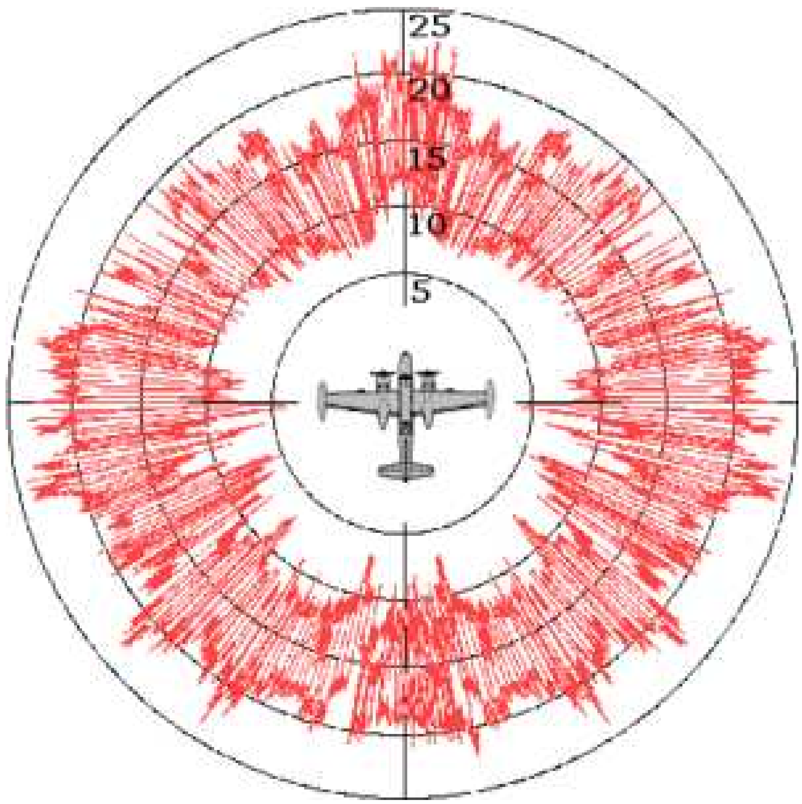


Figure 46: Typical RCS Diagram (Wikipedia)

The simulation used in this dissertation for hypothesis testing is capable of incorporating data about the RCS signatures of aircraft of interest, including the directionality and bandwidth range for the signatures. These signature models interact with the function of the radar models in the FLAMES simulation. However, the ability to generate the anticipated signature of an aircraft is not inherent in FLAMES. Because of this limiting factor, and the desire to keep this dissertation unclassified, a generic survivability term is used in place of the RCS of each vehicle of interest. This term was based on publicly available information and is meant to show only that incorporation of survivability information is possible with the methodology presented in this dissertation and the direction of the impact on the capability.

Table 12: Survivability Factor

System	Survivability Factor
B-1B	1
B-52	100
F-111	33
TLAM	0

4.4.3.3.2 Electronic Warfare

During a military strike mission, the aircraft delivering the ordinance is rarely the only aircraft involved in the mission execution. In addition to the strike platform there are often airborne control aircraft, refueling assets, and electronic warfare aircraft. These electronic warfare assets are important to increasing the probability that a mission will be successful by increasing the survivability of the strike platform. The DoD defines EW as "military action involving the use of electromagnetic and directed energy to control the electromagnetic spectrum or to attack the enemy. Electronic warfare consists of three divisions: electronic attack, electronic protection, and electronic warfare support." [129]

Electronic attack is defined by the same source as the "division of electronic warfare involving the use of electromagnetic energy, directed energy, or anti-radiation weapons to attack personnel, facilities, or equipment with the intent of degrading, neutralizing, or destroying enemy combat capability and is considered a form of fires."

According to the DoD, electronic protection is the "division of electronic warfare involving actions taken to protect personnel, facilities, and equipment from any effects of friendly or enemy use of the electromagnetic spectrum that degrade, neutralize, or destroy friendly combat capability." [129]

Electronic warfare support is the "division of electronic warfare involving actions tasked by, or under direct control of, an operational commander to search for, intercept, identify, and locate or localize sources of intentional and unintentional radiated electromagnetic energy for the purpose of immediate threat recognition, targeting, planning and conduct of future operations." [129]

A common form of electronic warfare that impacts the survivability of aircraft is radar jamming. The DoD identifies two primary types of jamming under the broad umbrella of electromagnetic jamming: barrage jamming and spot jamming. Jamming is the use of electromagnetic energy to reduce the effectiveness of enemy electromagnetic capabilities. The difference between barrage jamming and spot jamming is that spot jamming is targeted at a specific frequency while barrage jamming encompasses a wide section of frequencies in which electromagnetic systems operate. [129]

If jamming technologies were to be incorporated into a modeling and simulation environment, the primary impact would be on the effectiveness of radar sensors for the adversary. This could be captured by modeling the physics of the electromagnetic interference or by a modification to the system level performance of the adversary radar. This could take the form of a generic reduction in sensor range, or a reduced likelihood of acquisition when under the effect of a jamming system. Jamming technology and tactics are closely guarded to reduce the effectiveness of the adversary's countermeasures, as are the changes in capability that they allow. In order to eliminate the possibility of inadvertently exposing sensitive material about the performance of systems in a jammed environment, electronic warfare of any type will not be considered in the modeling portions of this dissertation.

4.4.3.4 Modeling Munitions

According to the General Accounting Office, and contrary to popular belief, during Operation Desert Storm, "95 percent of the total bombs delivered against strategic targets were unguided; 5 percent were guided." The reasons for the use of so many unguided munitions included the high cost of guided bombs, the all-weather capability of unguided bombs, and the large size of many of the strategic targets. [97] Because of the large number of unguided munitions delivered, the modeling and simulation environment will focus on modeling only unguided munitions, specifically freefall bombs.

At the time of Operation Desert Storm, the B-1B was not configured to deliver conventional ordinance. [241]

4.4.3.5 Measures of Performance

Three measures of performance were tracked from the FLAMES parametric scenario generator for the three aircraft in the simulation. These measures were the number of blue aircraft shot down, the number of red targets destroyed, and the number of SAMs fired. The measures were tracked over 25 repetitions of the same scenario settings in order to account for the probabilistic nature of some of the variables in the simulation. While 25 cases would normally not be considered a large enough sample to gain a high-confidence in the outcomes, the need to limit computational expenditure, coupled with the conceptual nature of the problem, led to the selection of this number of cases.

The measures of performance tracked in the scenario allow the quantification of the mission success rate (how many red targets were killed), the friendly loss rate (how many blue aircraft were shot down), and a rough measure of cost to the enemy through the number of SAMs fired. The cost to the US forces is dependent on whether the blue bomber was shot down or survived. The cost model used for these scenarios is discussed in the following section.

4.4.3.6 Costs

Before discussing the costs used in this experiment, it is important to note that these costs are not meant to make actual military decisions, but rather show the ability of regret analysis to incorporate cost information. Assumptions have been made where data was unavailable that may lead to results that differ from reality. In each case that such an assumption has been made, every effort has been made to clarify the reasoning behind the assumption.

The cost of a typical mission in Operation Desert Storm for each of the aircraft listed above was taken from General Accounting Office documents. Those costs, coupled with the unit replacement cost for each aircraft, are listed in Table 13. Figures for the average mission cost of the B-1B and B-52 were not available for Operation Desert Storm (the B-1B did not actually serve because of its nuclear armament in 1991). The operation cost for the B-52 was estimated based on the cost of the F-111. Based on the logic that was used to estimate the B-52 mission cost, the same approach was used to estimate the B-1B mission cost.

Table 13: Aircraft Costs

Aircraft	Mission Cost	Cost of Crew	Unit Cost (FY98)	Total Cost of Loss
	$C_{mission}$	$C_{crew} = \#_{crew}(\$6M)$	C_{unit}	$C_{mission} + C_{unit} + C_{crew}$
B-1B	$50,000	$24,000,000	$283,100,000	$307,150,000
B-52	$100,000	$36,000,000	$53,400,000	$89,500,000
F-111	$24,900	$12,000,000	$75,000,000	$87,024,900
TLAM	$2,855,000	N/A	N/A	N/A

The value of human life is something that is nearly impossible to quantify, and doing so opens a Pandora's Box of questions about ethics and morality. However, only labeling the loss of the aircraft as a financial burden is an incomplete assessment, especially when the aircraft being compared have different numbers in their crews. While either is terrible, the loss of a B-52, with its crew of six, will have a greater impact than the loss of the F-111, with its crew of two. In 1999, Conetta and Knight reported that "The Air Force estimates that it costs $6 million to train a pilot to full operational competence" [51] This figure, coupled with the unit cost of the aircraft in the mission, can give a sense of the cost to the US in the case of an aircraft being shot down.

4.4.4 Local Fitness Function

The fitness function used for the assessment of regret is shown in Equation 14. The aircraft fitness is the ratio of the probability of mission success to the expected cost of the mission. The Greek letters alpha, beta, gamma, and epsilon are weighting factors that can be used by decision makers to adjust assumptions (such as the cost for B-1B and B-52 missions) that they find objectionable and instantaneously see the effect on the analysis results. The factor on the crew cost, beta, can allow the incorporation of a higher value on human life than the training cost of $6 Million that was included in the initial experiment. Each cost in the fitness function is normalized by a baseline, shown in Table 14. These baseline values help condition the problem to reduce the likelihood of computational rounding errors.

$$Fitness = \frac{P_{mission\_success} * \alpha}{P_{aircraft\_lost} * \left(\frac{\beta * C_{crew}}{C_{crew\_baseline}} + \frac{\gamma * C_{unit}}{C_{unit\_baseline}} \right) + \frac{\varepsilon * C_{mission}}{C_{mission\_baseline}}}$$

Equation 14

Table 14: Baseline Cost Values

Cost Parameter	Baseline Value
Crew	$6,000,000
Unit	$50,000,000
Mission	$25,000

The local regret for each aircraft alternative is calculated based on the difference between the aircraft's fitness and the maximum fitness exhibited for that particular scenario. The aircraft that performs the "best" based on the fitness function therefore has a local regret value of zero. This formulation allows the regret function to be in a standard form for optimization or other computer manipulation. [243]

This form of local fitness function is often referred to in literature as an Overall Evaluation Criterion (OEC), which combines a number of metrics of interest into a single overall score for the particular concept. The challenge in using an OEC lies in the assignment of the weights for the various metrics. While analysts and engineers may establish weights, it is the decision makers who must have the final say in the importance of the various metrics. Therefore, the OEC creation must be an iterative process that involves the decision makers' input. This input can be solicited ahead of time, or, sometimes more effectively, solicited in an interactive electronic design review. This type of review allows instantaneous incorporation of decision maker feedback by leveraging the power of parametrics in early conceptual design.

4.4.5 Surrogate Model Creation

Because of the elaborate nature of the interactions in the FLAMES agent-based environment, the most likely candidate for a good surrogate model fit is the Artificial Neural Network (ANN). However, for completeness, two other forms of surrogate models will be considered, the polynomial response surface, and a krigining model. Because Global Regret is formulated as a series of integrals, the simplicity of the functional form will have great impact on the ability of the integrals to be calculated in closed form. While the Response Surface Methodology (RSM), ANNs, and kriging models are all functions that can be integrated, the complexity of the ANN and kriging functions can be limiting if more than a very few dimensions are considered.

4.4.5.1 RSM

The first attempts at fitting the measures of performance from the parametric scenario generator were based on the RSM. This method was chosen because it provides the simplest form of surrogate model, and is not very computationally intensive to create. The model created used a Response Surface Equation (RSE) form as shown in Equation 15 and is fit using a least-squares regression. In Equation 15, R is the response, the betas are the partial regression coefficients, and x's are the *n* regressors.

$$R = \beta_0 + \sum_{i=1}^{n} \beta_i x_i + \sum_{i=1}^{n} \beta_{ii} x_i^2 + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \beta_{ij} x_i x_j$$

Equation 15

Because of the discrete nature of the input for aircraft type, a separate surrogate had to be created for each aircraft. The results of the fitting are shown in Figure 47, Figure 48, and Figure 49. Because the TLAM's effectiveness is calculated based on an equation, no surrogate was created for it.

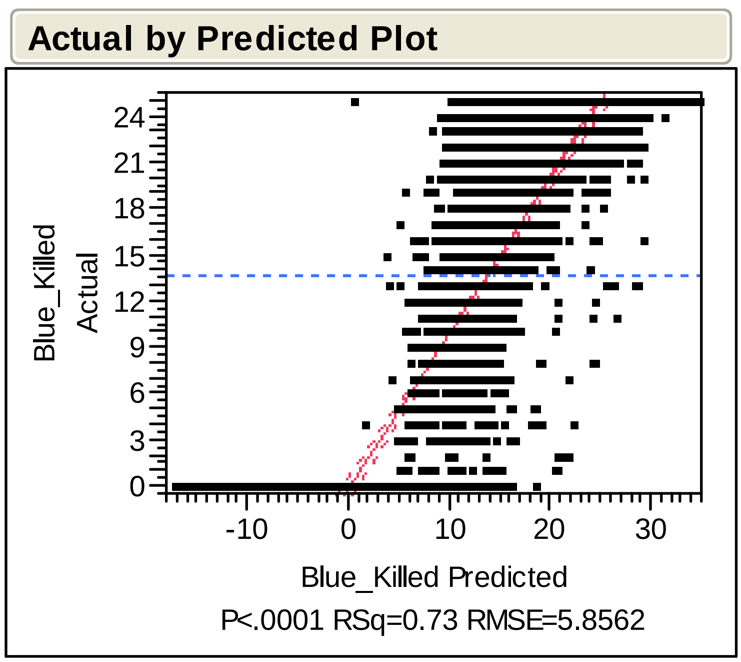


Figure 47: B-1B Units Lost RSE Fit

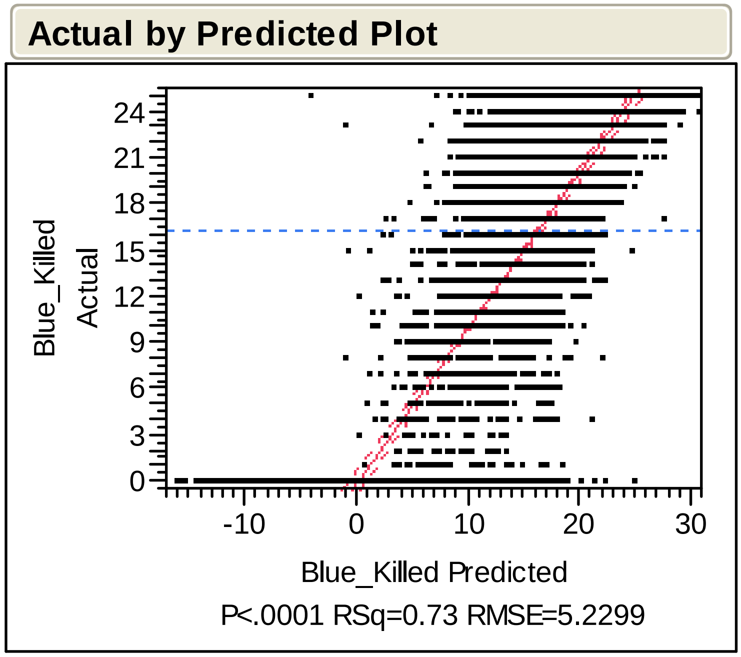


Figure 48: F-111 Units Lost RSE Fit

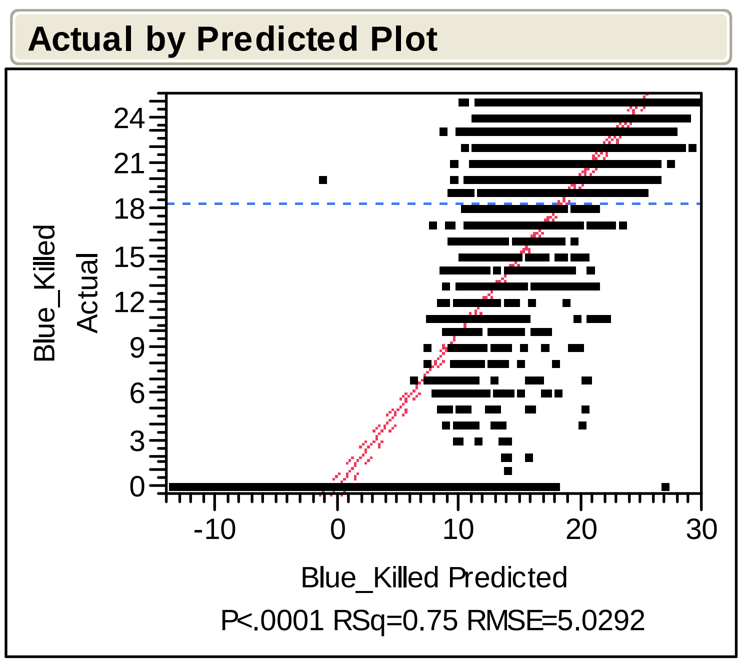


Figure 49: B-52 Units Lost RSE Fit

In the figures above, the number of blue bombers lost as predicted by the RSE is plotted on the horizontal axis, and the actual number lost according to the data is plotted on the vertical axis. The first measure of goodness of fit typically checked for surrogate model creation is the coefficient of determination, RSq, which indicates what portion of the response, on a scale of 0 to 1, is explained by the factors under consideration. If the RSq, were equal to 1, all points would lie along a 45 degree angle from (0, 0) to (25, 25). Because the RSq value for the responses is relatively low for engineering design standards, the RSE is considered a poor choice for describing the responses for these ranges in the parametric scenario generator. Biltgen typically found RSq values around 0.8-0.95 in his exploration of a three day strike scenario. [24]

4.4.5.2 ANN

Artificial Neural Network surrogates were built using the Neural Network tool in JMP 7.0 and also using BRAINN 2.0, a MATLAB ANN program created at the Aerospace Systems Design Laboratory (ASDL) at the Georgia Institute of Technology.

The JMP surrogates were trained using a random-holdback crossvalidation, with 25% of the data held back to check against overfitting. Because ANNs are capable of accounting for discrete variables, only one ANN was required for each of the responses of interest. The best fit obtained with the JMP neural network fitting tool was an RSq of approximately 89%, found with 13 hidden nodes Figure 50. The performance limitations of the JMP tool allowed only a range of 3-14 hidden nodes to be explored during the initial ANN fitting experiment.

Fit History

Nodes	Penalty	RSquare	CV RSquare
3	0.001	0.80891	0.81233
5	0.001	0.84440	0.84434
7	0.001	0.85995	0.86373
9	0.001	0.00000	-297.90
9	0.001	0.87168	0.87429
13	0.001	0.88341	0.88621
13	0.001	0.89060	0.89096
14	0.001	0.84378	0.84678
12	0.001	0.87280	0.87525

Figure 50: Fit History for Blue Bombers Lost

Figure 51 shows the fit results obtained using the BRAINN software to build the surrogate model of the number of blue bombers lost. Because of the automated nature of the BRAINN software, a wide range of numbers of hidden nodes could be explored relatively quickly. The BRAINN software was used to explore fits for configurations of ten to forty hidden nodes, using three iterations at each setting, in intervals of five hidden nodes, using a Gradient Descent with Moment Adaptive Learning Rate because of the large number of cases used in fitting. This approach is suggested by Johnson and Schutte in the BRAINN 2.0 manual. [127] A setting of 30 hidden nodes was found to provide the best fit for the data, with an RSq of approximately 93%. Because of the similarity in the data, the sweep of possible ANN configurations was not conducted for the number of red targets destroyed. Rather, the same configuration of 30 hidden nodes was used, with an RSq of 91% (Figure 52).

Figure 51 and Figure 52 are each composed of four sub-charts. The top two sub-charts show histograms of the distribution of the Model Fit Error (MFE) and Model Representation Error (MRE). The MFE describes how well the ANN fits the data that was used to train it, while the MRE describe show well the model fits test data that was not included in the training process. In both figures, the errors are centered about 0 with a standard deviation of less than 1, which meets the generally accepted criteria for errors surrogate model building. [168] There is some concern caused by the pattern in the Residual by Predicted plot, which displays a pattern on the edges, but given the highly conceptual nature of the selection problem, and the wide range of conditions over which the Measures of Effectiveness were considered, this is somewhat expected.

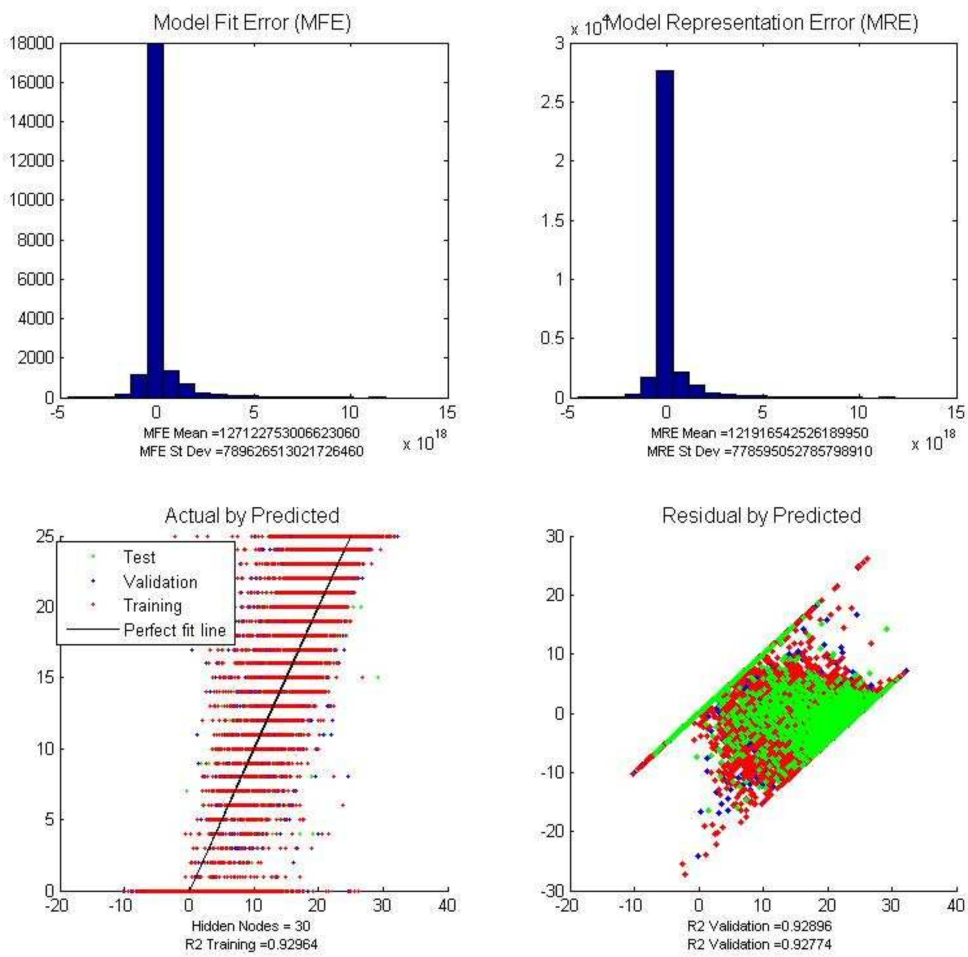


Figure 51: BRAINN Fit Results for Blue Bombers Lost

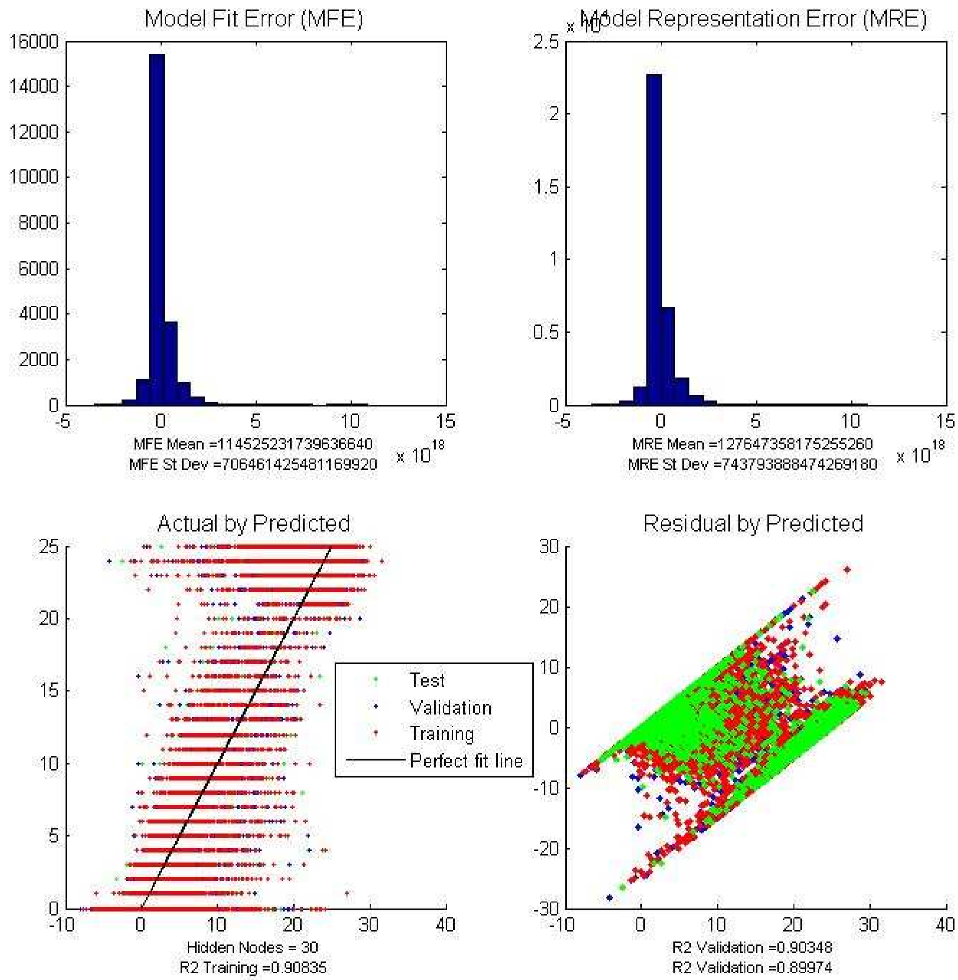


Figure 52: BRAINN Fit Results for Red Targets Destroyed

The 91-93% fits found using BRAINN were a few percentage points improvement over the JMP neural network tool, and a significant improvement over the polynomial response surfaces. However, because the number of nodes found to be appropriate using BRAINN was much larger than the number initially tried in JMP, an attempt was made to re-fit a 30 hidden node ANN in JMP. The results of this trial, and RSq of 91.086% for the blue aircraft lost, did not show much of an improvement over the 13 hidden node ANN, and still fell short of the performance exhibited in BRAINN. As a result of these experiments, the ANN created with BRAINN will be used for the rest of the hypothesis testing experiments.

4.4.5.3 Kriging and Radial Basis Functions

While initially considered for use in this experiment, because of the large number of data points included in the design of experiments for the parametric scenario generator, kriging surrogates were not built. Fitting the kriging model would have required computational power beyond that available for this dissertation.

4.4.6 Sub-Hypothesis Discussion

Based on the fit results from the surrogate model experiments, it is true that surrogate models can be fit to metrics across a wide array of possible futures. These surrogate models can be used anywhere in the possible future space defined by the ranges on the scenario variables.

The use of the surrogate models for Global Regret Analysis also addresses one of the research objectives. The flexibility and parametric nature of the analysis are important because of the trend toward electronic design reviews in the defense acquisition process. [156], [13] These electronic design reviews exist to incorporate decision maker feedback rapidly into the design process. Analysts running the codes are usually not qualified to establish the values of the elements in the objective function for the candidate evaluation (Equation 14). Any discrepancy between the values assigned to the objective function and those held by the decision makers has the potential to discredit the analysis. However, if the analysis is done in such a way that it can be rapidly assessed, such is the case when using surrogate models, the values of the weightings in the objective function can be changed on-the-fly.

The parametric nature of the design review allows credibility to be gained through tuning of assumptions to those acceptable by the decision makers; however, the volume of data can be overwhelming. In these cases, sensitivity studies can be carried out beforehand to understand which assumptions and design/scenario variables actually have a significant impact on the results. By only presenting the dimensions of the problem that have a significant impact, the understanding can be gained without overwhelming the decision makers. It is important, however, to keep the full data set available to defend any questions about the validity of the sensitivity studies.

4.4.7 Determination of Global Regret

With surrogate models created for the MOEs for each platform, a number of options exist for the evaluation of Global Regret. Under the formal definition of Global Regret, presented in Equation 9, the Global Regret would be calculated in its closed form, by integrating the ANN equations over each of the 10 dimensions of the parametric scenario generators. This integration can either be approached analytically, numerically, or another approach can be used to evaluate the Global Regret.

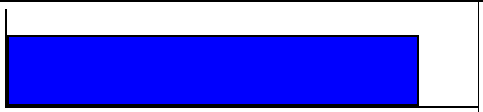
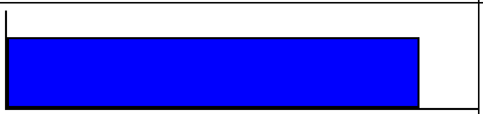
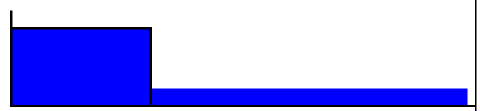
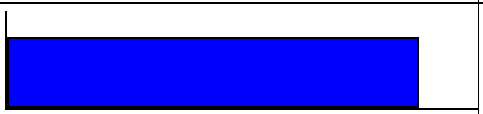
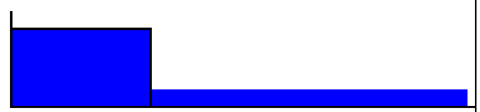
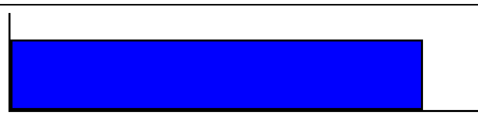
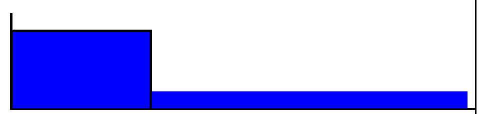
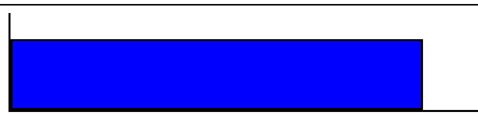
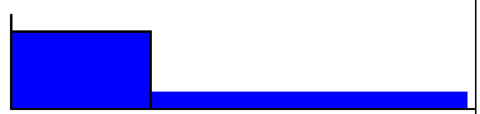
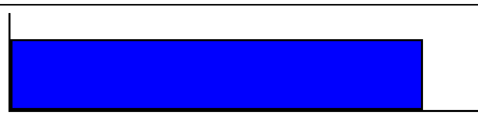
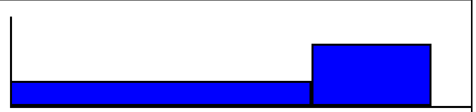
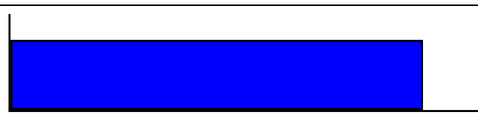
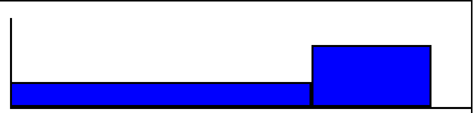
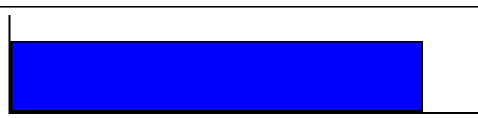
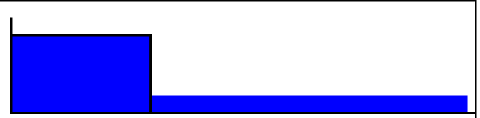
The integration over 10 dimensions provides a significant challenge. Because the number of evaluations required for integrating numerically grow exponentially with each additional dimension and the "curse of dimensionality" is quickly encountered. Monte Carlo methods exist that can help overcome the necessity of this numerical integration. In this experiment, a Monte Carlo simulation will be used, and the statistical distribution of the results will be used to understand the trends in the Global Regret for each of the candidates. While this does not result in the true Global Regret as defined mathematically, it provides a significant amount of information and can be used as an approximation for the Global Regret.

4.4.8 Scenario Probabilities

Because the scenario framework of interest for this experiment is Operation Desert Storm, an effort will be made to increase the likelihood of parameters that reflect the capabilities of the Iraqi Military in that period. As was defined earlier, the likelihood a particular scenario is the joint probability of all the parameters in the scenario. These parameter probabilities must be set by the decision makers, based on their understanding of the likely progression of military technology. Because of this dependence on expert judgment, the results contained herein are meant to be representative of a process only. Actual numbers should not be considered.

The likelihood functions for each of the eight adversary scenario variables are shown below in Table 15, where the x-axis is the range of the scenario variables, and the y-axis is the likelihood of a particular scenario. In the actual experiment, two weightings were used, one with uniform probabilities on the ranges and one reflective of Desert Storm parameters. The stepped distribution represents a probability of four times more for any scenario in the higher-probability region than those in the low-probability region. This allows comparison of the robustness of concepts that show little knowledge of the likelihood of futures a priori with those the robustness of those that only consider a smaller set of possible futures. Because altitude and speed are controlled by mission planners, these will be allowed to vary over their full ranges to find the most suitable flight condition for the particular scenario.

Table 15: Scenario Variable Probabilities

Scenario Variable	Distribution 1	Distribution 2
Distance from SAM Site		
SAM Pk		 Step occurs at Pk = 0.5
SAM Thrust		 Step occurs at Thrust = 750 lbs
SAM Max G		 Step occurs at Max G = 12
Ground Radar Transmission Power		 Step occurs at Power = 45 dBW
Ground Radar Scan Period		 Step occurs at Scan period = 20 sec
Ground Radar SNR Threshold		 Step occurs at SNR Threshold = 13
Ground Radar Acquisition Range		 Step occurs at Range = 200 km

4.4.9 Case 1: F-111 versus B-52

This case explores the set of scenarios that would exist if a conflict circa 1991 were to have sea access denied and, consequently, no TLAM presence in theater. This might be the case had operations taken place in Afghanistan during the early 1990's.

4.4.9.1 Distribution 1

The first investigation allowed the scenarios to vary with equal likelihood for each possible future scenario; each variable was assigned a uniform distribution for 4000 cases. Four thousand cases were used because they filled the design space adequately, but still provided good responsiveness for the JMP software. The JMP software tended to slow down significantly with more than 5 – 6 thousand cases, as each point is re-drawn with any changes to the scatterplots.

Figure 53and Figure 54 show the impact of the various factors on the OEC for the F-111 and B-52, respectively. In both cases, the distance the aircraft was able to maintain from the SAM site had the greatest impact on the cost weighted likelihood of success of the mission. Aircraft altitude and speed, both parameters that are controllable by mission planners and the pilots, were also important factors for both cases.

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
SAM Dist	0.0663777	0.002229	29.78	<.0001*
Altitude	-4.957e-5	6.642e-6	-7.46	<.0001*
Speed	-0.001202	0.000658	-1.83	0.0676
SAM_Thrust	-0.000439	0.000266	-1.65	0.0989
SAM_PK	-0.551307	0.439554	-1.25	0.2098
SAM_MaxG	0.0066954	0.011034	0.61	0.5440
GR_SNRThreshold	-0.009026	0.01754	-0.51	0.6068
GR_AcquisitionRange	1.8827e-7	5.246e-7	0.36	0.7197
GR_ScanPeriod	0.0029595	0.010887	0.27	0.7858
GR_TransmissionPower	-0.000657	0.013141	-0.05	0.9601

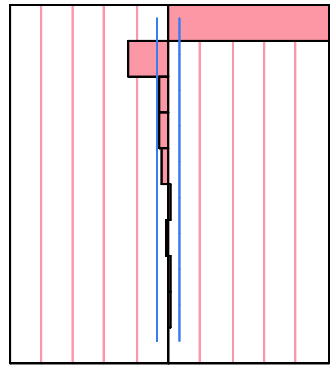


Figure 53: F-111 OEC Contributing Factors

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
SAM Dist	0.0197518	0.000579	34.11	<.0001*
GR_TransmissionPower	-0.030011	0.003414	-8.79	<.0001*
Altitude	-1.427e-5	1.726e-6	-8.27	<.0001*
GR_SNRThreshold	0.0242593	0.004557	5.32	<.0001*
Speed	-0.00071	0.000171	-4.16	<.0001*
SAM_PK	-0.216603	0.114208	-1.90	0.0580
GR_ScanPeriod	0.0012381	0.002829	0.44	0.6616
SAM_MaxG	0.0008073	0.002867	0.28	0.7783
SAM_Thrust	1.6459e-5	6.913e-5	0.24	0.8118
GR_AcquisitionRange	2.2474e-8	1.363e-7	0.16	0.8691

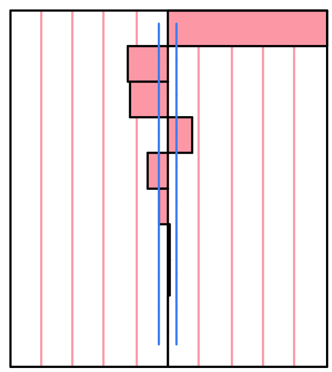


Figure 54: B-52 OEC Contributing Factors

Interestingly, the performance parameters for the F-111 fell into three groups by order of importance: mission planning variables, SAM variables, and finally radar parameters. Because the F-111 was traveling at relatively high Mach numbers for a large number of the mission cases, the most significant factor in the loss of blue aircraft appears to be whether the SAM was able to reach the F-111 in time. In the case of the B-52, traveling at subsonic speeds, the SAM site had the opportunity to launch multiple times, making the performance of the SAM less important relative to some of the radar parameters.

4.4.9.2 Sub-Hypothesis Testing

An extremely useful tool in the JMP software allows the visualization of 3-D plots that will update in real-time as other variables are changed. Figure 55 through Figure 58 show two examples, one for the F-111 and one for the B-52, of the changes that can be seen by using the 3-D surface plots and changing other variables.

In Figure 55 and Figure 56, the F-111 OEC is plotted against the ground radar SNR threshold, the most significant of the ground radar parameters, and the SAM probability of kill, the second most important SAM parameter. The difference between the two figures is the altitude at which the F-111 is flying, with 30,000 feet occurring in the first figure and 40,000 feet in the second. As can be seen from the shape of the surface in the figures, at the lower altitude, the coupling between the radar parameter and SAM parameter is not apparent. In the first figure, the F-111 OEC decreases with an increasing SAM probability of kill (the SAM is more likely to destroy the aircraft), and decreases slightly with the SNR threshold (the radar is more likely to differentiate the aircraft from ambient noise at a lower SNR threshold). However, at 40,000 feet, the relationship is slightly more complex. The overall OEC is lower at 40,000 feet, with a maximum of approximately one quarter the magnitude of the aircraft flying at 30,000 feet. Additionally, the SAM probability of kill does not impact the OEC above 0.5 for any value of SNR threshold. However, there is a coupling between the SNR threshold and the SAM probability of kill that is visible in the left side of the plot. In this region, the aircraft has a higher OEC, which can be explained by the low SAM performance and the low radar performance. In this region the radar does not detect the F-111 until it is closer, and has only a limited opportunity to shoot at the aircraft. If the aircraft is missed on the first shot (or the second) it is unlikely that the F-111 will still be within the detection range of the radar. Above a SAM probability of kill of 0.5, however, only a few shots are needed to down the blue aircraft, reducing its OEC to zero.

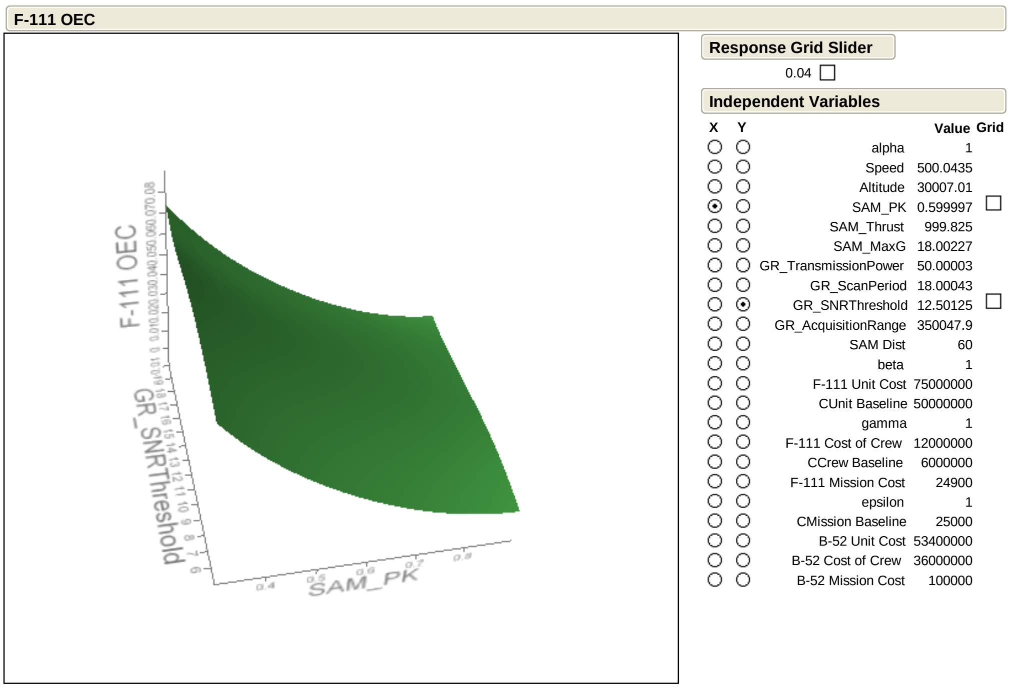


Figure 55: F-111 Performance, 30k ft

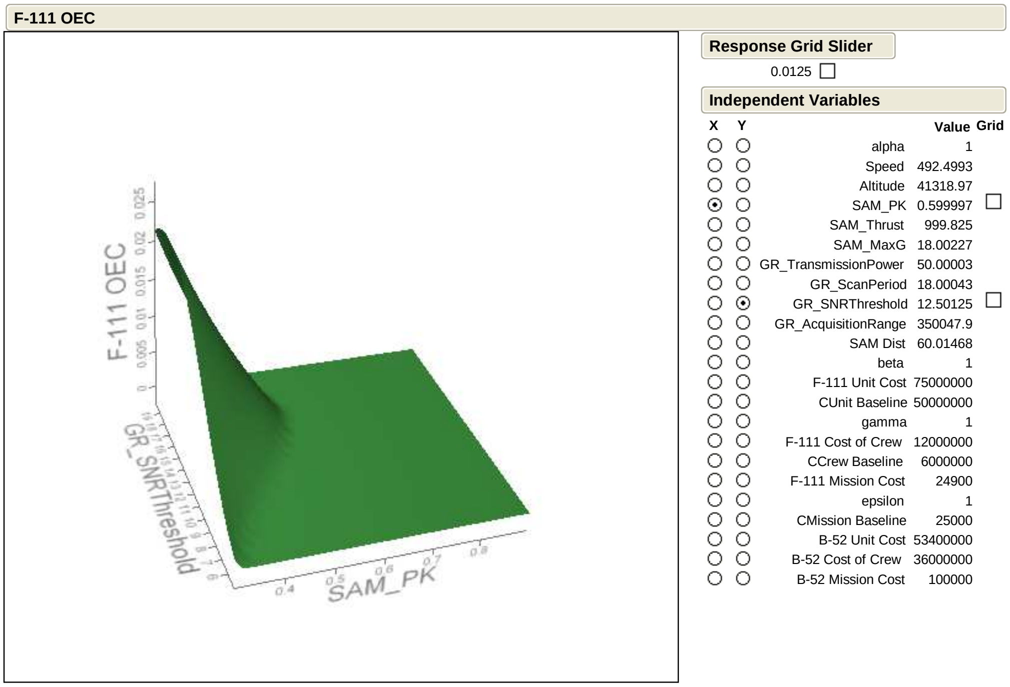


Figure 56: F-111 Performance, 40k ft

Figure 57 and Figure 58 show the OEC for the B-52 as a function of the radar transmission power and the SAM probability of kill. The difference between the two plots is the distance by which the B-52 is able to avoid the SAM site, with the first figure showing a 60 km distance and the second showing a 90 km distance. The surface plot in Figure 57 shows that the primary driver on the B-52's OEC is the SAM probability of kill, which, as it improves, lowers the B-52 OEC. This trend is consistent with the expected outcome of a SAM of increasing capability against a fixed platform. This trend is consistent across the range of ground radar transmission powers explored.

At a longer distance from the SAM site, however, the dominant trend is reversed. The OEC magnitude is significantly greater at the longer distance (which is expected because of the reduced threat environment), but the primary driver on the OEC is the radar transmission power. As the transmission power increases, the radar site is able to detect the B-52 at longer ranges, and has the opportunity to shoot more SAMs. There is a slight increase in the B-52 OEC with the decrease in SAM probability of kill at lower radar powers, but because the SAM site is able to shoot so many times at the higher powers, there is little effect in that region.

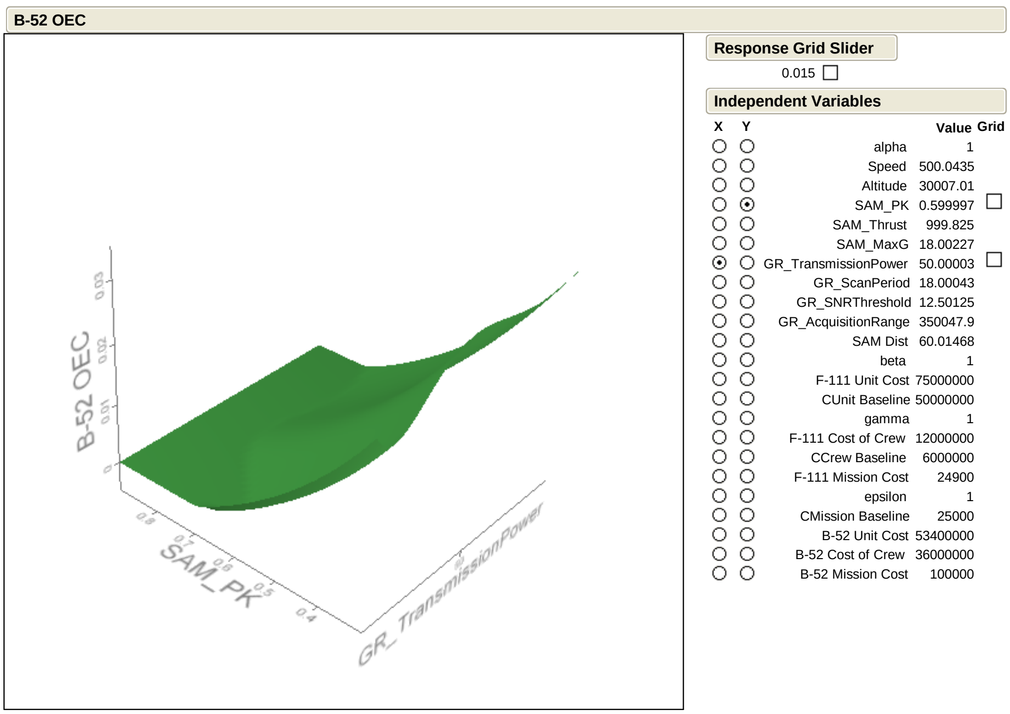


Figure 57: B-52 Performance, 60 km keep-out

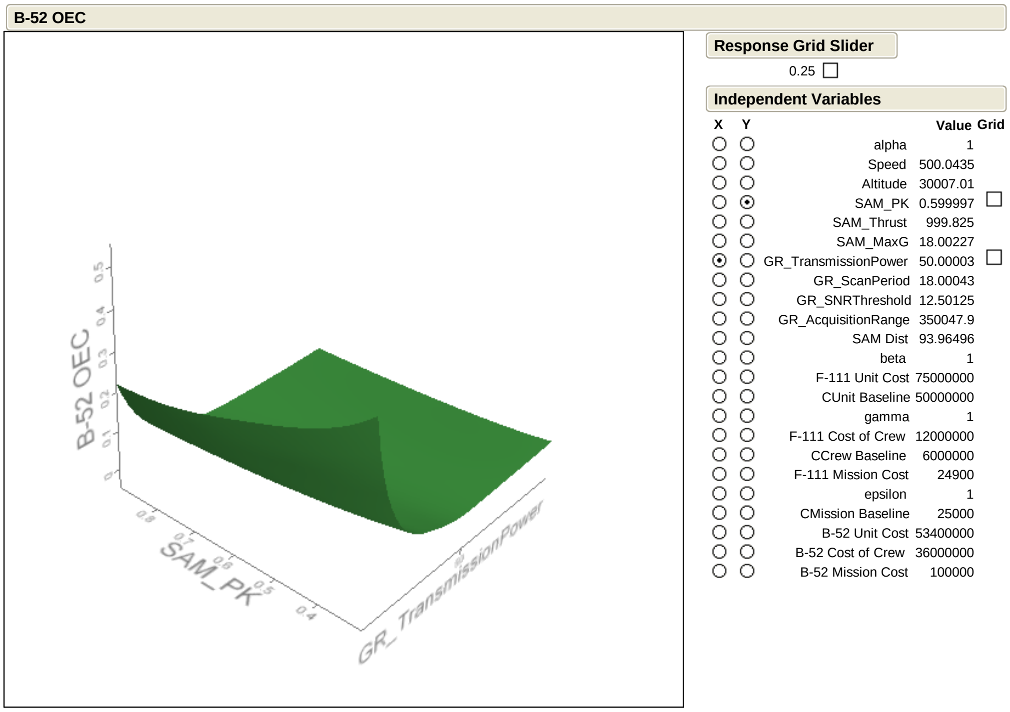


Figure 58: B-52 Performance, 90 km keep-out

4.4.9.3 Sub-Hypothesis Conclusions

These two examples have shown the use of surface plots to explore the behavior of systems across many possible scenarios. Without the use of surrogate models, which allow the partial differentials to be understood, the creation of these plots would not be possible, and electronic design reviews could not rapidly explore the alternative's effectiveness. This demonstrates the ability of surrogate models to enable interactive, electronic design reviews using possible futures.

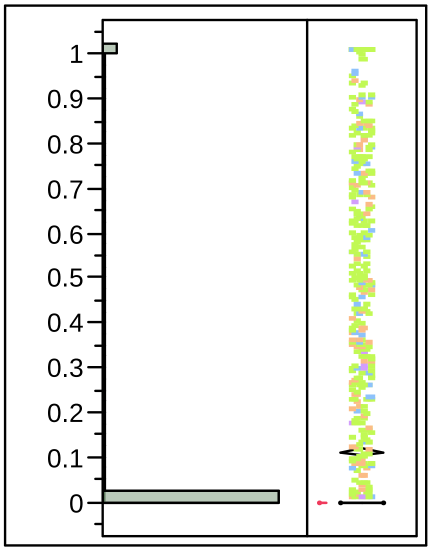
4.4.9.4 Evaluating Regret for Case 1

Evaluating the regret across the entire scenario space can be accomplished by either evaluating the integral form of the Global Regret Equation, shown in Equation 9, or by examining the statistical parameters associated with the evaluation of the randomly chosen cases across the scenario space. Because of the complexity of the integral form, for this case statistical data will be used to evaluate the Global Regret and decide which candidate is more suitable for the missions based on the OEC. It is important to note that the weightings in the OEC would be tuned in an interactive, electronic design review, so the results shown here are to demonstrate the method only, not suggest Air Force policy.

Figure 59 shows the statistical data that was created for the regret of the F-111 and B-52 using the statistical package JMP. The candidate alternative with the best performance will exhibit zero local regret for a particular scenario, so smaller is better for the statistics. This side by side comparison shows that both candidates have at least 25% of cases where they exhibit zero regret. This implies that there is no “magic bullet solution” between the two aircraft. However, at least 75% of the scenarios showed the F-111 as a better candidate based on the OEC. Also, the mean regret for the F-111 is approximately 25% of the mean regret for the B-52. Based on these statistics, the Global Regret for the F-111 is lower than the Global Regret of the B-52 and would be the more robust choice for the set of possible futures explored in this experiment.

Distributions

F-111 Local Regret



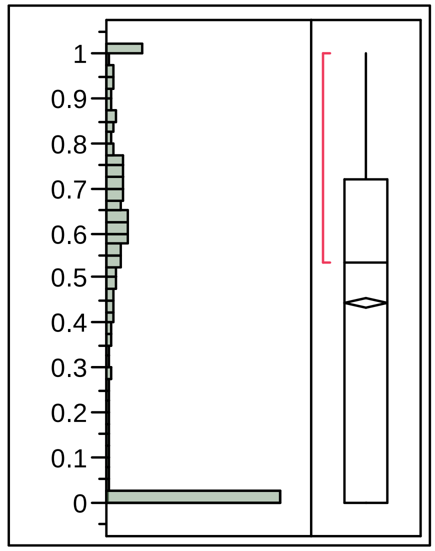
Quantiles

100.0%	maximum	1.0000
99.5%		1.0000
97.5%		1.0000
90.0%		0.6184
75.0%	quartile	0.0000
50.0%	median	0.0000
25.0%	quartile	0.0000
10.0%		0.0000
2.5%		0.0000
0.5%		0.0000
0.0%	minimum	0.0000

Moments

Mean	0.1110576
Std Dev	0.2881152
Std Err Mean	0.0045555
upper 95% Mean	0.1199889
lower 95% Mean	0.1021263
N	4000

B-52 Local Regret



Quantiles

100.0%	maximum	1.0000
99.5%		1.0000
97.5%		1.0000
90.0%		0.9143
75.0%	quartile	0.7207
50.0%	median	0.5334
25.0%	quartile	0.0000
10.0%		0.0000
2.5%		0.0000
0.5%		0.0000
0.0%	minimum	0.0000

Moments

Mean	0.4439169
Std Dev	0.351082
Std Err Mean	0.0055511
upper 95% Mean	0.4548002
lower 95% Mean	0.4330337
N	4000

Figure 59: B-52, F-111 Regret Comparison

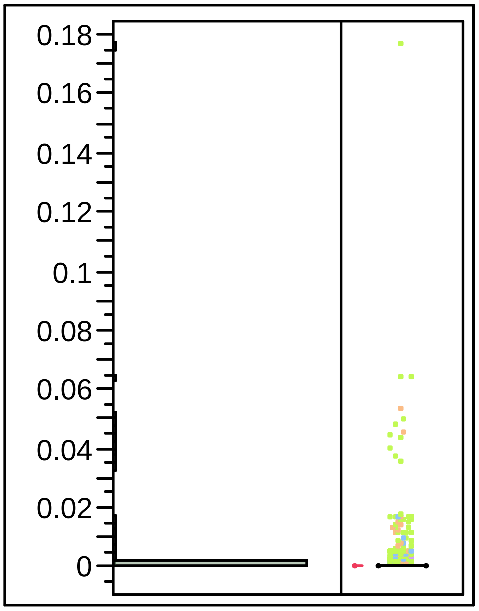
4.4.9.5 Comparison with Gulf War Operations

The same types of analyses shown in the sub-hypothesis testing can be used for the second set of distributions on scenario variables. However, because the trends are the primary interest for this dissertation, and the trends are not significantly impacted by the probability distributions, the explorations using the surface plots will not be repeated. However, the regret associated with two candidates does show a change when the probabilities consistent with the Gulf War threat environment are given higher likelihood than those in the rest of the scenario space.

Figure 60 shows the statistical data on the distribution of regret for the Gulf War threat environment. One immediate difference between the Gulf War set and the full, equally weighted data set is that the maximum regret for both the F-111 and B-52 are lower by 70% to 80%. This indicates that the candidates are much closer to each other in terms of performance across this region of the scenario space. However, the distributions of the areas of zero regret have also shifted. The 90th percentile regret for the F-111 is an order of magnitude lower than that of the B-52, whereas in the entire scenario space it was on the same order of magnitude (smaller by 30%). This indicates that over the Gulf War scenario space, for the mission type simulated, the F-111 was a better choice across the vast majority of possible scenarios. This is echoed in the mean regret for the two platforms, where the F-111 mean local regret is 30% of the B-52 mean local regret.

Distributions

F-111 Local Regret



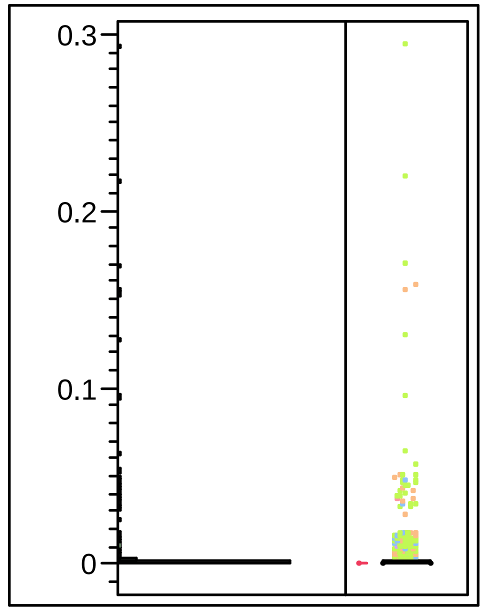
Quantiles

100.0%	maximum	0.17512
99.5%		0.01563
97.5%		0.00352
90.0%		0.00024
75.0%	quartile	0.00000
50.0%	median	0.00000
25.0%	quartile	0.00000
10.0%		0.00000
2.5%		0.00000
0.5%		0.00000
0.0%	minimum	0.00000

Moments

Mean	0.0004331
Std Dev	0.0039717
Std Err Mean	0.0000628
upper 95% Mean	0.0005562
low er 95% Mean	0.0003099
N	4000

B-52 Local Regret



Quantiles

100.0%	maximum	0.29319
99.5%		0.04196
97.5%		0.01088
90.0%		0.00270
75.0%	quartile	0.00066
50.0%	median	0.00014
25.0%	quartile	0.00000
10.0%		0.00000
2.5%		0.00000
0.5%		0.00000
0.0%	minimum	0.00000

Moments

Mean	0.00148
Std Dev	0.0086192
Std Err Mean	0.0001363
upper 95% Mean	0.0017472
low er 95% Mean	0.0012128
N	4000

Figure 60: B-52, F-111 Regret Comparison - Gulf War

Having determined through simulation that the F-111 would have been a better choice for most missions in Operation Desert Storm than the B-52, it is important to cross check this conclusion against data from the conflict. The General Accounting Office published the data in Table 16, which shows the number of total F-111 and B-52 strikes and the casualty rate per strike. Because the casualty rate is a significant factor in the OEC calculation, it would be expected that the F-111 would be used in more strikes, have a lower casualty rate, or both. This is in fact the case, with the F-111 flying approximately 50% more missions than the B-52 and having a casualty rate of approximately one-third that of the B-52. This corroboration of historical events with the simulation data strengthens both the model accuracy and the method appropriateness for this type of problem.

Table 16: GAO Aircraft Casualty Rates in Operation Desert Storm [97]

Aircraft	Total Casualties	Total Strikes	Aircraft Casualty Rate per Strike
F-117	0	1,788	0
F-111F	3	2,802	0.0011
F-15E	2	2,124	0.0009
A-6E	8	2,617	0.0031
O/A-10	20	8,640	0.0023
F-16	7	11,698	0.0006
F/A-18	10	4,551	0.0022
B-52	5	1,706	0.0029
GR-1	10	1,317	0.0076

4.4.10 Case 2: Introduction of TLAM Technology

This case explores the scenario most representative of the actual Operation Desert Storm. Three major systems that were used for the majority of the bombing campaign are compared with this set of possible friendly alternatives. Figure 61 through Figure 63 show the influence of the various scenario variables on the local regret for each system. With the introduction of the TLAM, the tactics are no longer the most important variables for determining which system is more appropriate for the mission. The tactics are both important, but radar parameters, and SAM parameters are also important. However, the most important factor is still the distance by which the aircraft can evade the SAM site.

The change from the OEC to the local regret as the metric of interest for comparing concepts is important with the introduction of the TLAM. Because the TLAM was build on a strictly probabilistic model based on TLAM performance in the Gulf War, the OEC for the TLAM is fixed. This essentially establishes a baseline that shows how well, or poorly, the two other alternatives are doing relative to the cruise missile. Because of this status as a baseline, the local regret factors for the TLAM represent the factors that are the most significant in driving the TLAM to a regret value above zero. In that case, the tactics variable altitude does play a significant role in the regret. Following the tactics and SAM distance, the two factors that affect how quickly the radar can detect the aircraft, transmission power and SNR threshold are most important. The final factor in the group of “heavy hitter” variables is the SAM probability of kill, which relates to how many shots the SAM site has to be able to take to down the aircraft.

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
SAM Dist	0.0384815	0.004866	7.91	<.0001*
SAM_Thrust	-0.001091	0.000581	-1.88	0.0605
GR_TransmissionPow er	-0.032035	0.028692	-1.12	0.2643
Altitude	-1.534e-5	1.45e-5	-1.06	0.2904
SAM_PK	0.7998122	0.959728	0.83	0.4047
GR_AcquisitionRange	-8.191e-7	1.146e-6	-0.72	0.4746
GR_SNRThreshold	0.0133711	0.038296	0.35	0.7270
SAM_MaxG	-0.007996	0.024091	-0.33	0.7400
Speed	-0.000118	0.001436	-0.08	0.9342
GR_ScanPeriod	0.0015388	0.02377	0.06	0.9484

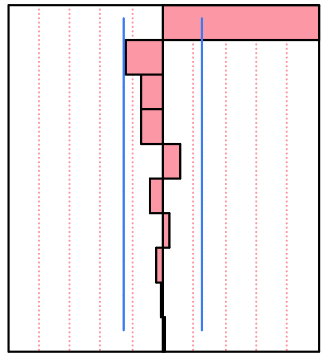


Figure 61: F-111 Local Regret Factors

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
SAM Dist	0.0920331	0.034848	2.64	0.0083*
GR_AcquisitionRange	1.3476e-5	8.203e-6	1.64	0.1005
GR_ScanPeriod	-0.260837	0.170219	-1.53	0.1255
SAM_PK	-10.48577	6.872621	-1.53	0.1272
Altitude	-0.000155	0.000104	-1.50	0.1347
Speed	-0.013992	0.010281	-1.36	0.1736
SAM_Thrust	0.0035057	0.00416	0.84	0.3994
SAM_MaxG	-0.07363	0.172515	-0.43	0.6695
GR_SNRThreshold	0.0935823	0.27424	0.34	0.7329
GR_TransmissionPow er	0.0378693	0.205465	0.18	0.8538

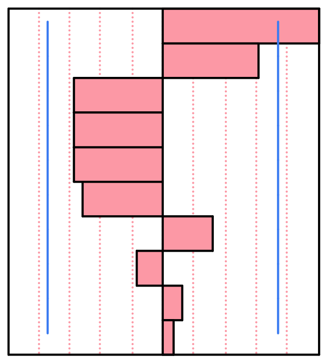


Figure 62: B-52 Local Regret Factors

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
SAM Dist	0.0049575	0.000113	43.97	0.0000*
Altitude	-2.118e-6	3.36e-7	-6.30	<.0001*
GR_SNRThreshold	0.0041986	0.000887	4.73	<.0001*
GR_TransmissionPow er	-0.002536	0.000665	-3.81	0.0001*
SAM_PK	-0.049748	0.022235	-2.24	0.0253*
GR_AcquisitionRange	3.1516e-8	2.654e-8	1.19	0.2351
GR_ScanPeriod	0.0005394	0.000551	0.98	0.3274
Speed	2.1475e-5	3.326e-5	0.65	0.5186
SAM_MaxG	-0.000286	0.000558	-0.51	0.6079
SAM_Thrust	6.8064e-6	1.346e-5	0.51	0.6131

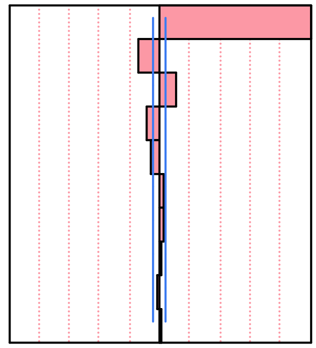


Figure 63: TLAM Local Regret Factors

The following figures (Figure 64, Figure 65, and Figure 66) show surface plots of the local regret versus the distance the SAM site is avoided by and the ground radar scan period. The ground radar scan period was used because it does not have a significant impact of the regret and allows the trends in the distance to be seen more clearly. In these cases, the most desirable state is where regret is zero, so when the surface is flat and at zero, that candidate is the "best" for that scenario. By looking at the values of local regret for each candidate as a function of the distance, the regions where each is best can be easily identified. At distances of less that 100 km, the TLAM shows zero regret, while the F-111 and B-52 both have positive regret values, indicating that the TLAM would be the best choice for this type of mission. This is consistent with the use of cruise missiles against targets in defended areas. Beyond 100 km, there is a region, which is still within striking distance of the most capable SAMs, where the F-111 is the best choice for the mission. Beyond 110 km, however, the B-52 becomes the most favorable platform. This is consistent with the differences expected between a strategic bomber and a strike aircraft, with the strike aircraft being used in riskier situations and the strategic bomber reserved for lower-risk engagements.

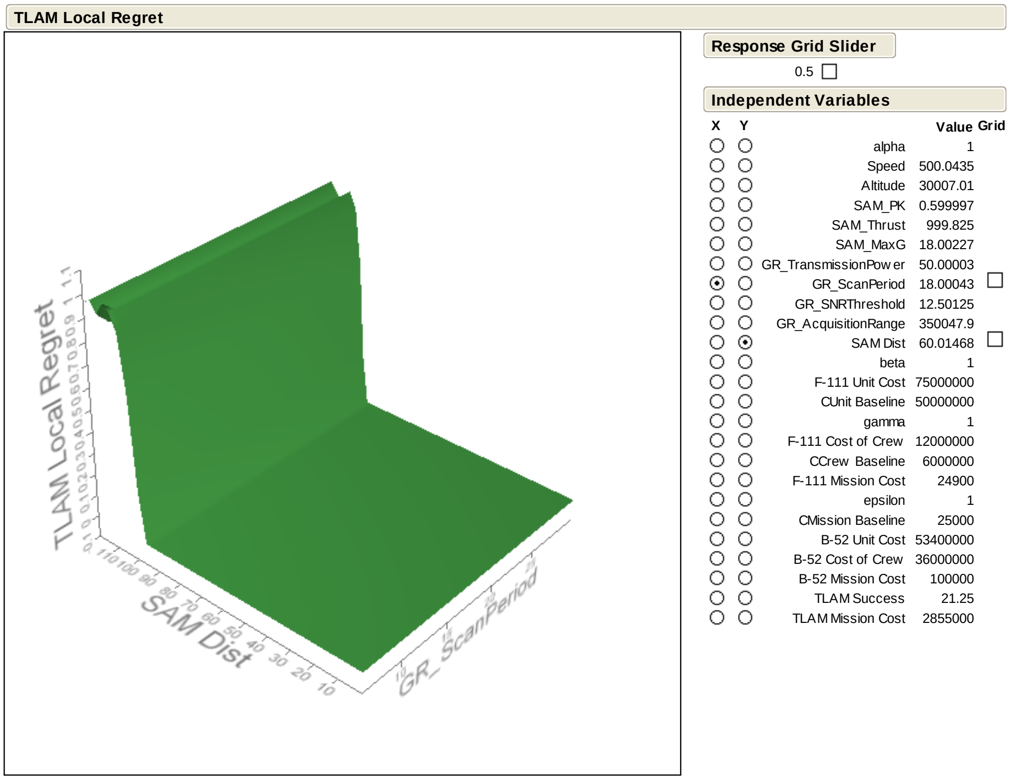


Figure 64: TLAM Regret vs SAM Distance and Radar Scan Period

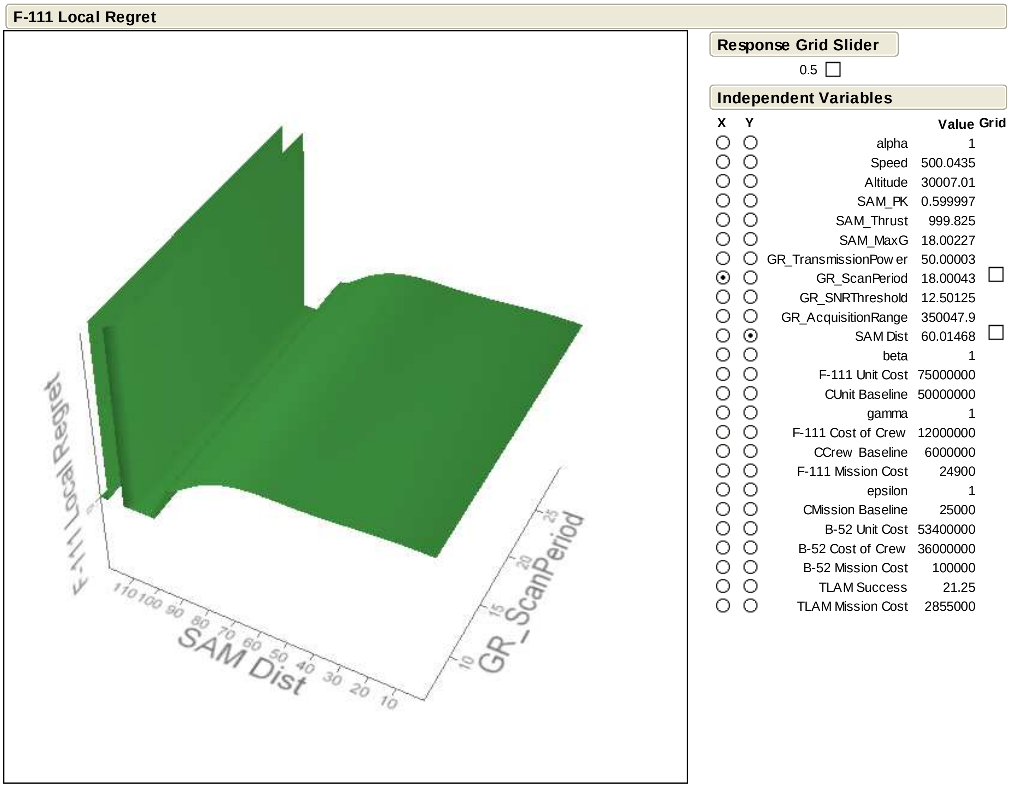


Figure 65: F-111 Regret vs SAM Distance and Radar Scan Period

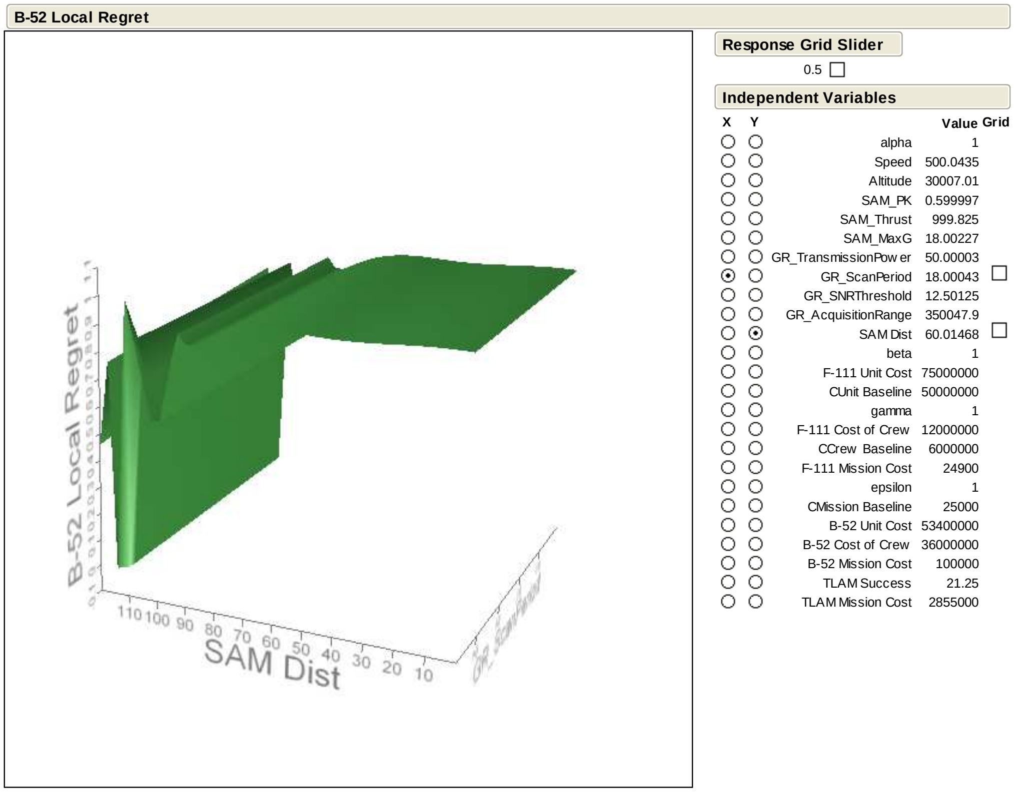
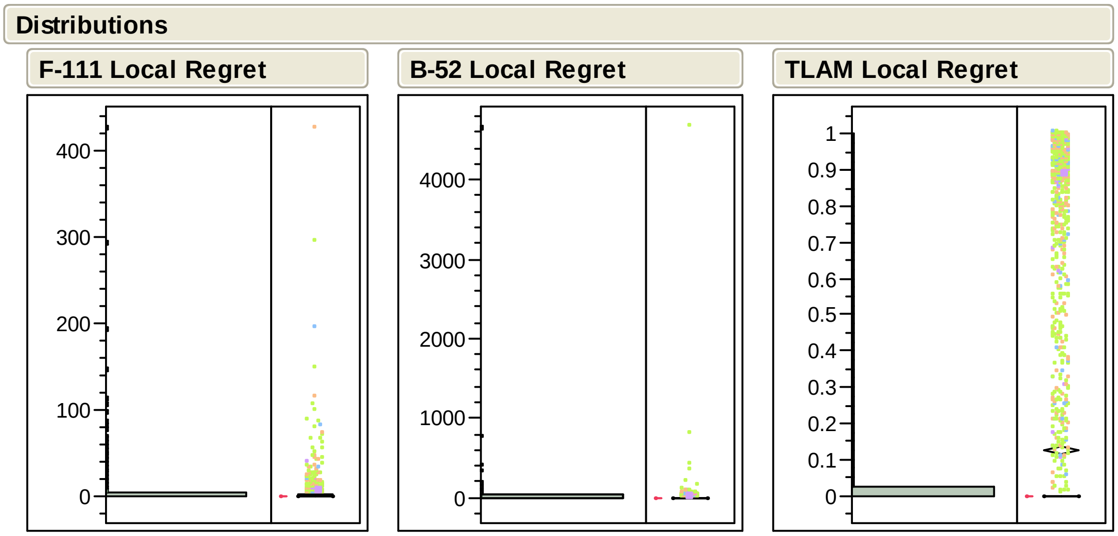


Figure 66: B-52 Regret vs SAM Distance and Radar Scan Period

Figure 67 shows the statistical data for the local regret for the three systems considered for this case. Looking first at the mean local regret as a measure of the Global Regret, the TLAM clearly has a significantly lower mean local regret than either the F-111 or the B-52. The F-111 is still a better choice than the B-52, which is consistent with the comparison done in the previous case. In fact, the TLAM does not show significant regret until the 90th percentile, while the F-111 shows significant regret in the 25th percentile and the B-52 in the 10th percentile. These results are consistent with the sizes of the zero regret areas in Figure 64 through Figure 66.



F-111 Local Regret

Quantiles

100.0%	maximum	424.62
99.5%		45.43
97.5%		8.86
90.0%		1.11
75.0%	quartile	1.01
50.0%	median	0.87
25.0%	quartile	0.59
10.0%		0.00
2.5%		0.00
0.5%		0.00
0.0%	minimum	0.00

Moments

Mean	1.725563
Std Dev	10.556223
Std Err Mean	0.1669085
upper 95% Mean	2.0527968
low er 95% Mean	1.3983292
N	4000

B-52 Local Regret

Quantiles

100.0%	maximum	4645.4
99.5%		30.3
97.5%		5.7
90.0%		1.0
75.0%	quartile	1.0
50.0%	median	0.94913
25.0%	quartile	0.82325
10.0%		0.46927
2.5%		0.0
0.5%		0.0
0.0%	minimum	0.0

Moments

Mean	2.9097788
Std Dev	75.103922
Std Err Mean	1.1874973
upper 95% Mean	5.2379353
low er 95% Mean	0.5816222
N	4000

TLAM Local Regret

Quantiles

100.0%	maximum	0.99975
99.5%		0.99464
97.5%		0.96608
90.0%		0.80181
75.0%	quartile	0.00000
50.0%	median	0.00000
25.0%	quartile	0.00000
10.0%		0.00000
2.5%		0.00000
0.5%		0.00000
0.0%	minimum	0.00000

Moments

Mean	0.1245196
Std Dev	0.2975454
Std Err Mean	0.0047046
upper 95% Mean	0.1337432
low er 95% Mean	0.1152959
N	4000

Figure 67: B-52, F-111, TLAM Regret Comparison

This comparison has shown the ability of the regret analysis approach to include a new technology that may be evaluated via a different evaluation criterion than the conventional alternatives. The new technology in this case performed better than the conventional alternatives in many possible scenarios, but there were scenarios where each of the other alternatives was more successful. Additionally, the analysis allowed understanding of the boundaries of best performance for each of the alternatives.

4.4.10.1 Comparison with Gulf War Operations

Because this set of alternatives is most representative of the aircraft available during Operation Desert Storm, the comparison with historical data has the most potential for supporting the use of the Parametric Scenario Generator coupled with Global Regret Analysis. The TLAM, B-52, and F-111 were all used extensively during the conflict.

When the probabilities of the Gulf War-like scenarios were increased, an interesting result occurred in the significant factors for the local regret of the aircraft and TLAM. The impacts of the various scenario variable factors are shown in Figure 68. While over the entire scenario space the tactics of the blue aircraft were more important than threat parameters associated with the radar or SAM, when emphasis is placed on the region of the scenario space closer to Iraq's capabilities, the tactics are least important. In this range of the space, range from the SAM site is still the dominant factor, but a set of three radar parameters are the next most important, followed by the SAM probability of kill. There is also a shift in the relative importance of the other factors. While in previous experiments, the distance from the SAM site was by far the most significant factor, in this case the importance of the following four factors is close to that of the distance.

One slightly counter-intuitive factor is the relationship between speed and regret. In this region of the scenario space, the speed is primarily impacting the accuracy of the bombing run, not the ability of the aircraft to leave the SAM's reach quickly. As speed increases, the accuracy of the bombing run decreases, decreasing the likelihood of a successful mission. Because this factor is the primary driver in the numerator of the OEC, it impacts the regret; as mission effectiveness goes down (due to missed bombing runs) regret goes up.

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
SAM Dist	0.0000372	5.189e-6	7.17	<.0001*
GR_TransmissionPow er	-0.000165	3.059e-5	-5.41	<.0001*
GR_ScanPeriod	0.0001265	2.535e-5	4.99	<.0001*
GR_SNRThreshold	0.0001819	4.083e-5	4.45	<.0001*
SAM_PK	-0.004191	0.001023	-4.10	<.0001*
GR_AcquisitionRange	-4.566e-9	1.221e-9	-3.74	0.0002*
SAM_MaxG	-9.084e-5	2.569e-5	-3.54	0.0004*
SAM_Thrust	-2.129e-6	6.194e-7	-3.44	0.0006*
Speed	0.0000028	1.531e-6	1.83	0.0675
Altitude	-2.119e-8	1.546e-8	-1.37	0.1707

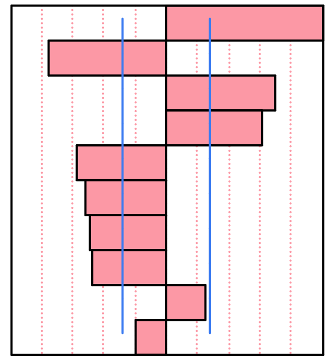
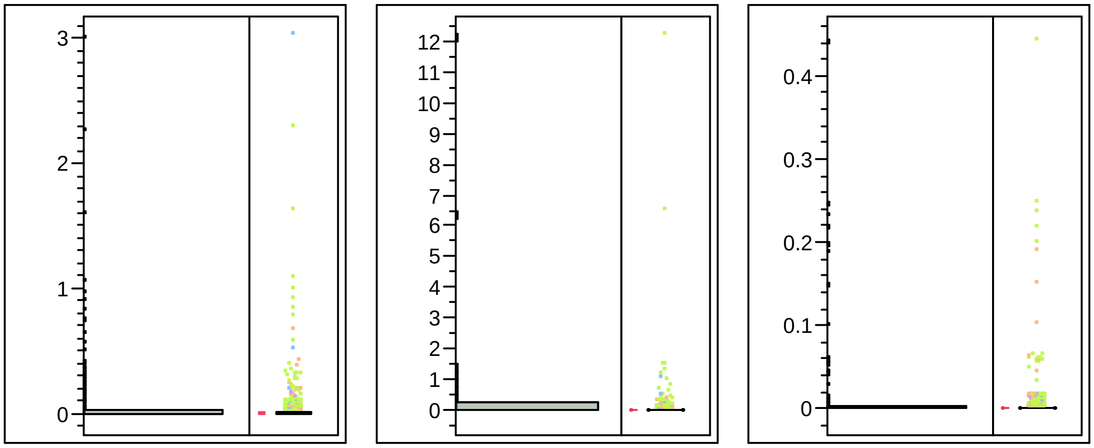


Figure 68: Local Regret Significant Factors - Gulf War Scenario

The statistics for local regret for the Operation Desert Storm case are shown in Figure 69. The trends among the three candidates are similar to those for the entire scenario space, but the magnitude of the regret associated with each alternative is significantly smaller. However, by weighting the scenarios the difference between the F-111 and B-52 has become smaller. While in the overall scenario space the F-111's mean local regret was 59% of the value for the B-52, in this weighting scheme the F-111's mean local regret is 67% of the B-52's. The change in the mean regret values relative to each other indicates that the alternatives are closer in OEC performance than they were over the entire scenario space.

Distributions

F-111 Local Regret	B-52 Local Regret	TLAM Local Regret



Quantiles

		F-111	B-52	TLAM
100.0%	maximum	3.0194	12.203	0.44243
99.5%		0.2859	0.255	0.05498
97.5%		0.0495	0.055	0.00695
90.0%		0.0099	0.013	0.00082
75.0%	quartile	0.0031	0.00364	0.00000
50.0%	median	0.0008	0.00093	0.00000
25.0%	quartile	0.00019	0.00024	0.00000
10.0%		0.0000	0.00006	0.00000
2.5%		0.0000	0.000	0.00000
0.5%		0.0000	0.000	0.00000
0.0%	minimum	0.0000	0.000	0.00000

Moments

	F-111	B-52	TLAM
Mean	0.0084903	0.0124031	0.0010833
Std Dev	0.077424	0.2248603	0.0114157
Std Err Mean	0.0012242	0.0035554	0.0001805
upper 95% Mean	0.0108904	0.0193735	0.0014371
low er 95% Mean	0.0060902	0.0054326	0.0007294
N	4000	4000	4000

Figure 69: B-52, F-111, TLAM Regret Comparison – Gulf War Scenario

4.4.11 Case 3: Conventional Arming of B-1B

This case explores what might have occurred if the B-1B conventional armament program had taken place before Operation Desert Storm. It allows the evaluation of a low-RCS candidate in the comparison among alternatives, and also gives some insight into the performance of the F-117, which has intentionally been omitted because many details of the F-117's performance, costs, and role in various conflicts remain classified.

The primary drivers for local regret for the three candidates are shown in Figure 70, Figure 71, and Figure 72. The F-111 and B-52 show similar results to those of earlier tests, with some slight re-ordering of the importance of the factors. This re-ordering occurs because while a factor may not have a great impact on the OEC of the particular alternative, if it impacts another alternative's OEC significantly the regret value will change. As a result of this coupling, comparing the influences on the regret provides insights into not just the factors that significantly influence the performance of a single alternative, but captures the coupling effects of changing parameters on the performance gap between alternatives.

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
SAM Dist	0.0259299	0.003492	7.42	<.0001*
GR_TransmissionPow er	-0.04059	0.020591	-1.97	0.0488*
SAM_PK	1.26056	0.688745	1.83	0.0673
SAM_Thrust	-0.000542	0.000417	-1.30	0.1938
GR_ScanPeriod	0.013126	0.017059	0.77	0.4417
Altitude	-6.349e-6	1.041e-5	-0.61	0.5419
GR_SNRThreshold	0.0112519	0.027483	0.41	0.6823
Speed	-0.000224	0.00103	-0.22	0.8281
GR_AcquisitionRange	-1.039e-7	8.221e-7	-0.13	0.8995
SAM_MaxG	0.0020081	0.017289	0.12	0.9075

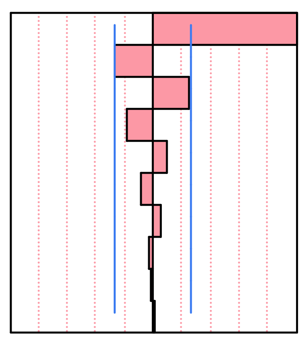


Figure 70: F-111 Local Regret Factors - Case 3

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
SAM Dist	0.0304746	0.007567	4.03	<.0001*
GR_SNRThreshold	0.1021102	0.059552	1.71	0.0865
GR_TransmissionPower	-0.062257	0.044618	-1.40	0.1630
SAM_PK	-2.01332	1.492421	-1.35	0.1774
Speed	-0.002834	0.002233	-1.27	0.2044
GR_ScanPeriod	-0.034304	0.036964	-0.93	0.3534
GR_AcquisitionRange	1.5768e-6	1.781e-6	0.89	0.3761
SAM_Thrust	-0.000724	0.000903	-0.80	0.4230
Altitude	-0.000017	2.255e-5	-0.76	0.4491
SAM_MaxG	-0.019995	0.037462	-0.53	0.5936

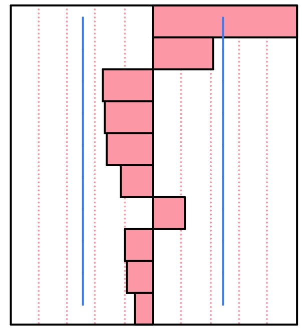


Figure 71: B-52 Local Regret Factors - Case 3

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
GR_TransmissionPower	-0.472495	0.193243	-2.45	0.0145*
GR_AcquisitionRange	-1.252e-5	7.715e-6	-1.62	0.1048
GR_SNRThreshold	0.4169518	0.257927	1.62	0.1061
SAM_Thrust	-0.005599	0.003912	-1.43	0.1525
SAM_PK	-7.110724	6.463803	-1.10	0.2714
Speed	-0.010294	0.00967	-1.06	0.2871
SAM Dist	0.0308568	0.032775	0.94	0.3465
GR_ScanPeriod	0.0503363	0.160093	0.31	0.7532
SAM_MaxG	-0.035443	0.162253	-0.22	0.8271
Altitude	0.0000152	9.768e-5	0.16	0.8764

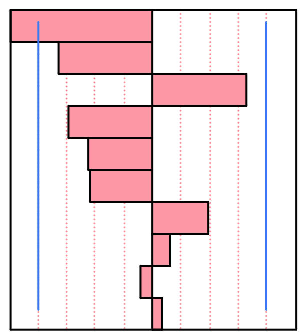


Figure 72: B-1B Local Regret Factors - Case 3

The local regret of the B-1B (Figure 72) is the first to have a set of major influences that is not topped by the distance from the SAM site. In fact, the distance is seventh on the list of the parameters. The primary drivers for the B-1B local regret are the three radar parameters that affect the detection of the aircraft, the parameter that affects how quickly the SAM can reach the aircraft, and how likely the SAM is to down the aircraft.

The importance of these factors makes sense in light of the RCS susceptibility parameter that was assigned to the B-1B. The factor makes the B-1B 100 times harder for the radar to detect than the B-52, and 33 times harder to detect than the F-111. By increasing the difficulty of detection for the radar, the range at which the radar does finally detect the B-1B will be significantly shorter. This short detection range means that the two SAM parameters relating to how quickly a SAM can engage the target, and how likely that first shot (which may be the only opportunity) is to destroy the aircraft are also very important.

The partial differentials of each of the variables for the B-1B can be seen in Figure 73. The regret for the B-1B increases for slower speeds, indicating that the SAM site is able to shoot multiple times. In situations where there are high losses, such as low speeds and at extremely close ranges to the SAM site, the regret for the B-1B is actually higher because of the very high aircraft cost. Cost is also the driving factor when the radar SNR threshold is very low. Because the radar can easily detect the aircraft, regardless of its susceptibility factor, many aircraft are shot down, making cheaper aircraft more desirable. It is important to note that the plots in Figure 73 are only valid at the settings of the scenario variables shown in red. When a variable changes, the plots update to reflect the new partial differential equation of regret for each of the candidate alternatives.

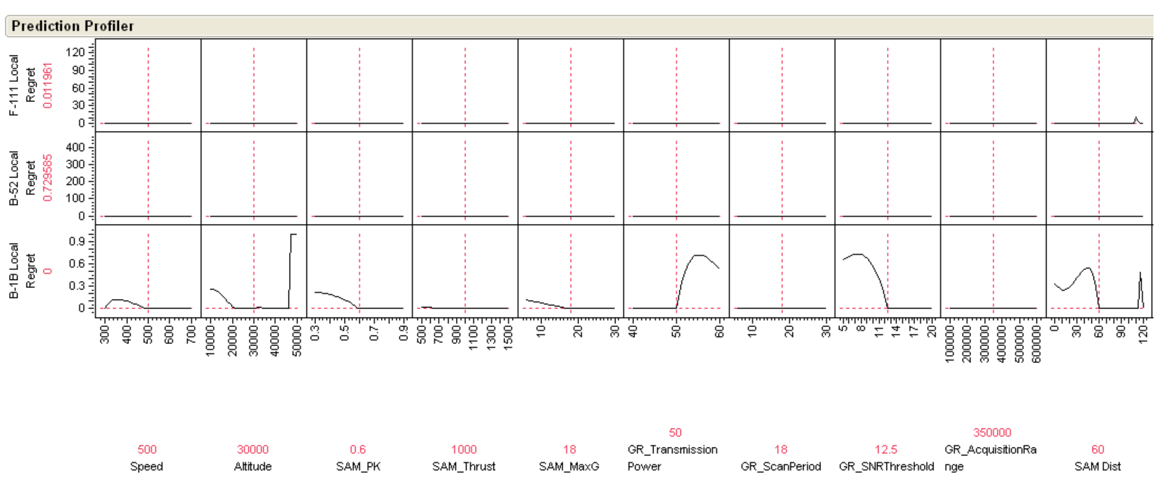


Figure 73: Regret Profiles for Case 3 – B-1B Best

The multivariate scatterplot for local regret shown is in Figure 74. In the scatterplot, the scenarios have been color coded by which alternative displays the best OEC: B-1B is best for gray points, B-52 is best for blue points, and F-111 is best for red points. This color coding allows regions of lowest regret to be identified in the scenario variables for each alternative, and also allows the regret for a particular candidate to be observed as a function of the candidate which is doing the best. The existence of the color trends identify regions where the parametric scenario generator can be used to gain insight into alternatives effectiveness in different regions of the design space.

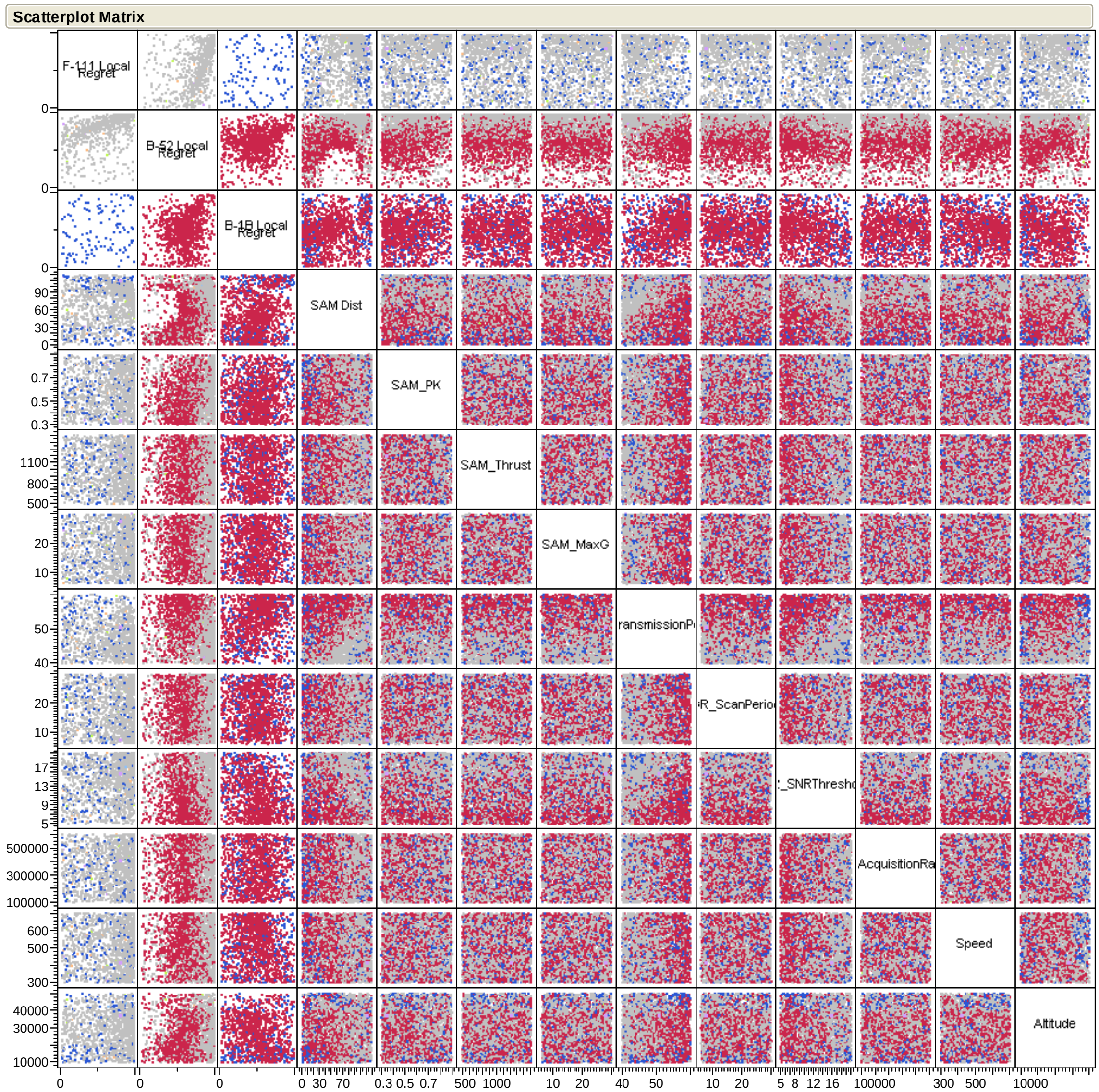


Figure 74: Regret Coloration Scatterplot

Figure 75 shows one example of three variables that definite trends appearing as color gradients. The trends show trades primarily between the F-111 (red) and B-1B (gray). In the top left scatterplot, SNR threshold is plotted against the radar transmission power. In the upper left corner of the plot, there are significantly more red points, and in the lower right there are significantly more gray points. This trend makes sense as the two variables affect the detection of the aircraft, and the F-111 is a significantly cheaper platform. In the upper left hand corner, it is much easier for the radar to detect and engage either aircraft, which results in more aircraft lost. In that case the cheaper aircraft is more attractive. However, in the lower right section of the plot, the radar is not as capable and has a much lower likelihood of detecting the B-1B. Because of this, even though the F-111 is a cheaper platform, because it is more likely to be shot down, the B-1B is the more attractive option. What this analysis basically boils down to is that when the radars are so powerful that the susceptibility of the aircraft isn't a factor, the cheaper aircraft is the better option.

An additional trend is visible in the top-right plot of Figure 75, though it is more subtle than the trend discussed above. In this plot, there is a region in the lower right that has fewer points where the F-111 is the best option, though it is not as distinct as the region in the top-left plot. This region corresponds to a high SAM probability of kill and a low radar transmission power. The reason that the B-1B is more desirable in this region is because the B-1B is more likely to be completely undetected by the radar for very low powers. However, when the SAM probability of kill is also low, the F-111 is likely to survive the attack, and therefore, as the cheaper platform, is more desirable.

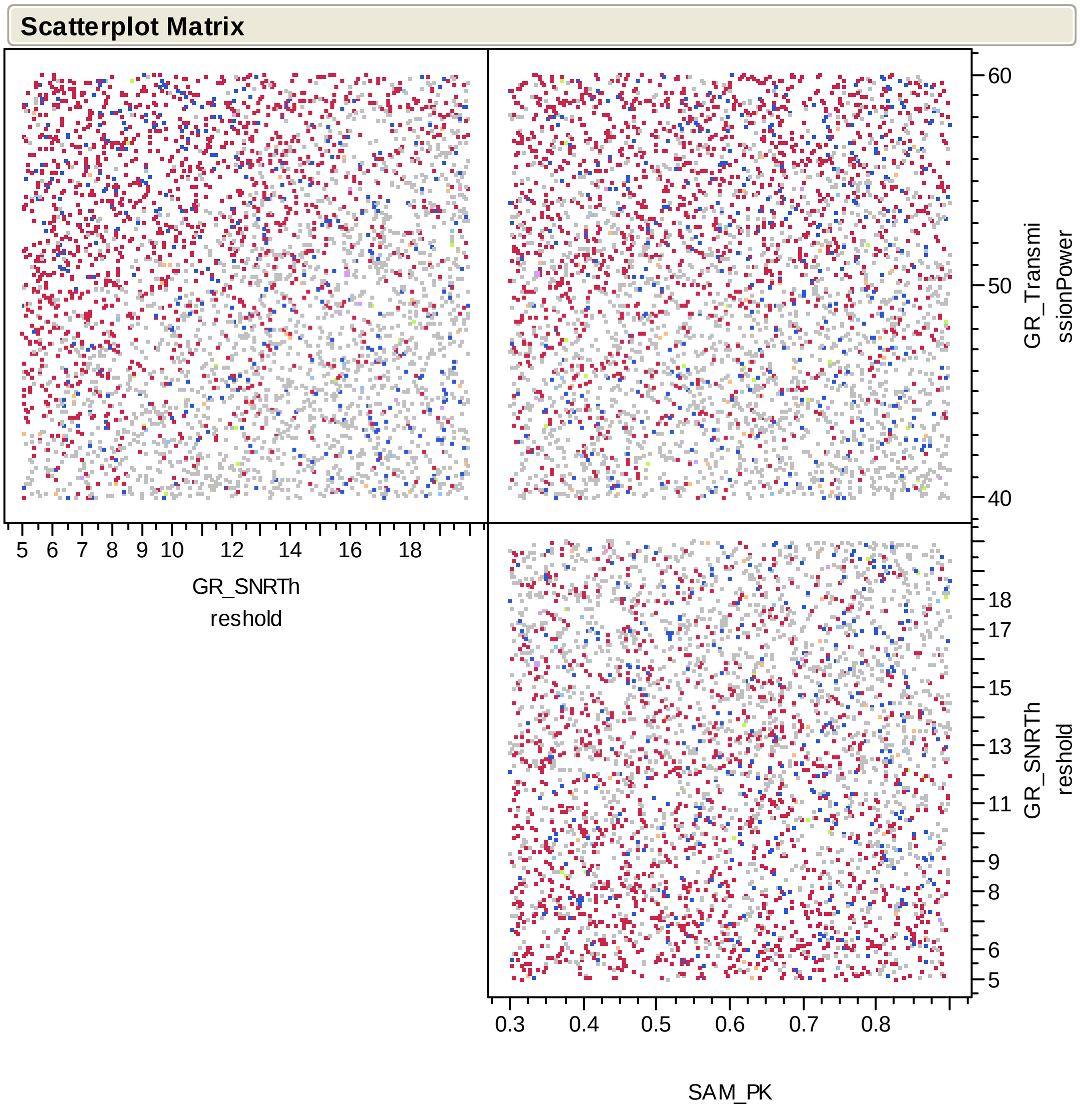
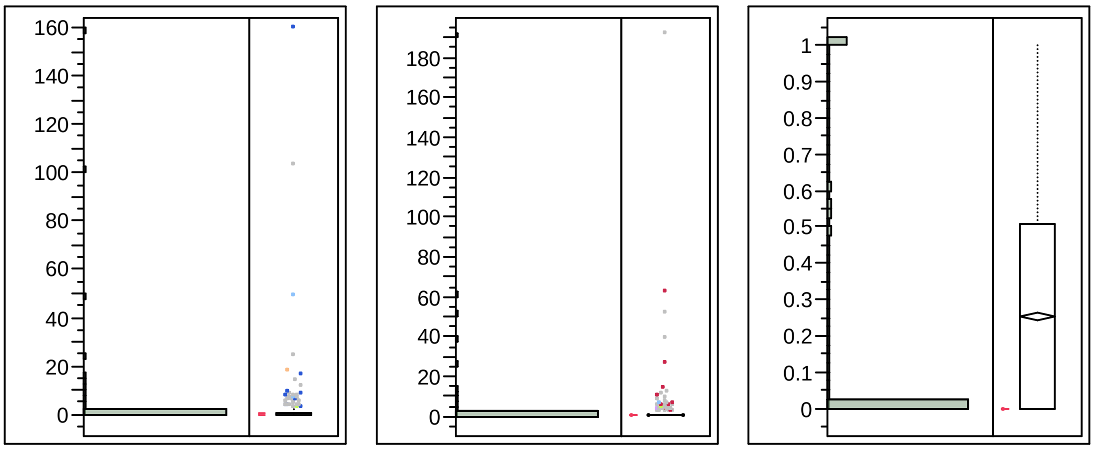


Figure 75: Three Key Variables for Regret

The scatterplot analysis suggests that the Global Regret assessment should identify the B-1B and the F-111 as the two most viable candidates. The statistical distributions of local regret for each of the three candidates are shown in Figure 76. Indeed, the mean local regret for the B-1B and F-111 are the lowest, with the B-1B having a mean of approximately 50% of the F-111 and 33% of the B-52. The standard deviation for the B-1B mean local regret is also extremely low, suggesting that for the majority of cases it is in fact the best option. The standard deviation of the B-52 and F-111 are similar, and the maximum local regret value for the two platforms is similar. This suggests, that even thought the F-111 was a better choice than the B-52, the difference between the two platforms is not as great as the difference between the B-1B and the F-111. This suggests that the B-1B is a significantly better choice than both the B-52 and the F-111.

Distributions

F-111 Local Regret	B-52 Local Regret	B-1B Local Regret



Quantiles

		F-111	B-52	B-1B
100.0%	maximum	159.00	191.10	1.0000
99.5%		5.27	4.71	1.0000
97.5%		1.41	1.37	1.0000
90.0%		1.00	1.00	0.8703
75.0%	quartile	0.97	0.99	0.5102
50.0%	median	0.35	0.74	0.0000
25.0%	quartile	0.00	0.40	0.0000
10.0%		0.00	0.00	0.0000
2.5%		0.00	0.00	0.0000
0.5%		0.00	0.00	0.0000
0.0%	minimum	0.00	0.00	0.0000

Moments

	F-111	B-52	B-1B
Mean	0.5891902	0.775594	0.2538
Std Dev	3.19126	3.3979866	0.3456534
Std Err Mean	0.0505468	0.0538212	0.0054748
upper 95% Mean	0.6882902	0.8811136	0.2645338
lower 95% Mean	0.4900903	0.6700744	0.2430662
N	3986	3986	3986

Figure 76: B-52, F-111, B-1B Regret Comparison

4.4.12 Summary of Experiment Results

The goal of this experiment was to use three cases to support the hypothesis that robustness could be defined as a function of the regret associated with alternatives. This hypothesis was supported by the experimental results. Global Regret Analysis successfully showed which candidates performed the best over a range of scenarios, the essential element of robustness.

Four primary methods were used for the visualization of the regret (and robustness) of alternatives: significant factors analysis, multivariate scatterplot analysis, partial differential analysis, and statistical analysis. The primary metrics identified for judging the robustness of a particular candidate are the mean local regret and the standard deviation of the local regret. The goal is to select a candidate with a mean local regret close to zero (preferably the lowest) and a low standard deviation. The low mean local regret means that over a large portion of the scenario space, the alternative was the best of those considered, with respect to the evaluation function. The low standard deviation indicates that in the regions where the candidate is not the best, it is not dominated by a significant margin.

Two cases of probability for scenarios were explored, the first where the regions of the scenario space were weighted equally and the second where the region representative of Operation Desert Storm was emphasized. When different regions were explored, different factors were the dominant influences on effectiveness and regret. Additionally, while the "best" candidate did not change when the smaller region was emphasized, the difference among the candidates shrank. The change in behavior across different regions of the scenario space underlines the need for design exploration beyond a small region of scenarios to understand how an alternative might truly perform in different types of conditions. As Lowry states, "the first thing to go after contact with the enemy is the plan.” [148] The uncertainty that surrounds warfare is enormous, and it is unlikely that a military systems will spend much time in the exact conditions for which it was designed. But by exploring the wide variety of possible futures, and exploring behavior within regions of those possible futures, better decisions can be made at the systems acquisition phase and also at the mission planning phase.

4.5 Experiment 3 – Filtered Monte Carlo Simulation w/ Possible Futures

4.5.1 Filtered Monte Carlo State-of-the-Art

Filtered Monte Carlo Simulation (FMCS) is a technique for performing top-down design based on a space composed of legitimate, discrete designs created in a bottom-up fashion. Thousands of individual designs spanning the design space are created, along with their metrics, and evaluated at each level of the SoS hierarchy. Each of these thousands of points is then a vector that contains a complete description of all the modeled aspects of the SoS and its performance metrics. Then constraints are applied at the top levels of the metrics to determine what range of alternatives has the potential of satisfying those constraints. [197], [78] This filtering can be done using a multivariate scatterplot, with areas allocated to each level of the SoS hierarchy as show in Figure 77.

Surrogate models enable FMCS techniques by allowing the rapid generation of thousands of cases for filtering. If a region of the design space is identified as promising, based on top-level filtering, thousands of additional points can be created instantaneously to explore trends in that area. This ability to rapidly "zoom-in" on a particular region of the design space is of particular use in interactive design reviews, as it allows the decision makers to understand trends in regions possibly not considered by analysts.

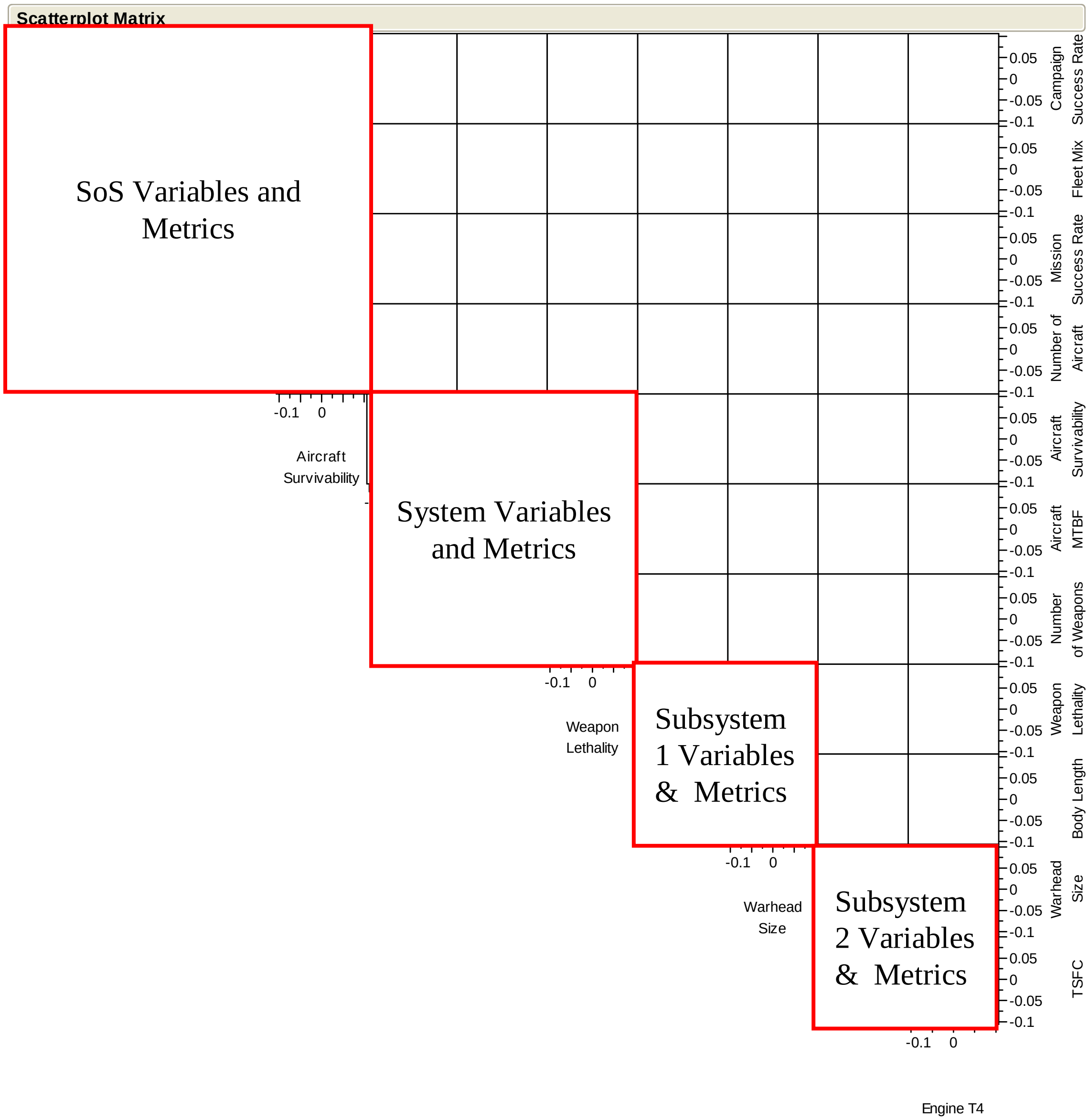


Figure 77: Basic FMCS Scatterplot

4.5.2 Expansion of Technique

The current state-of-the-art FMCS technique includes regions of space devoted to subsystem MOPs of performance, and SoS variables and MOEs. [24], [78] In order to incorporate information about possible futures space, additional space in the FMCS scatterplot will be allocated for scenario variables and local regret. The resulting FMCS scatterplot is similar to Figure 78, where the top-left corner of the scatterplot has been devoted to the scenario variables and local regret.

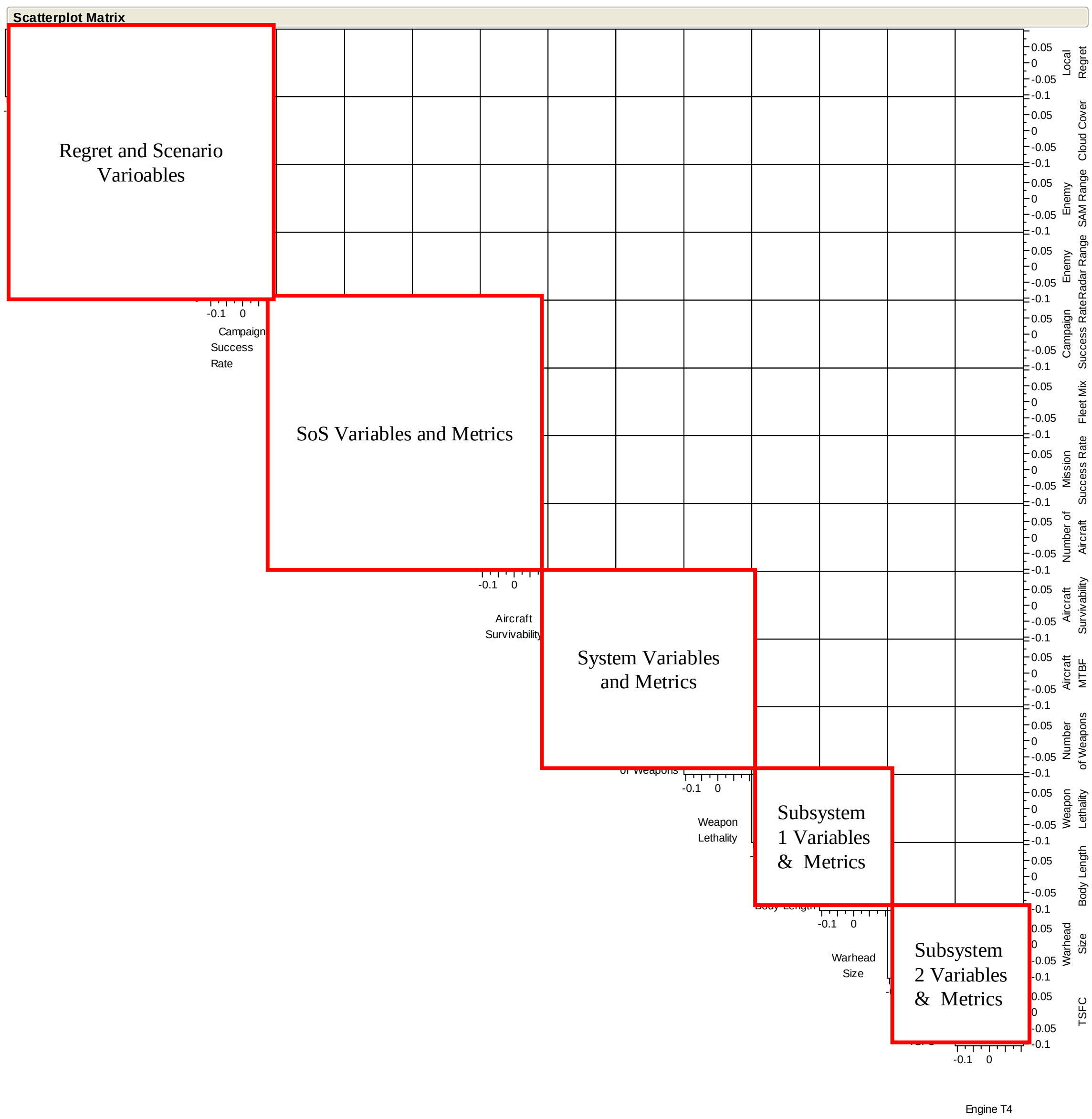


Figure 78: Expanded FMCS Scatterplot Structure

Including scenario variables and local regret in the FMCS scatterplot adds another dimension to the information available for decision making in electronic design reviews. While before top-level metrics were static, showing results only for the scenario or handful of scenarios for which they were created, now performance and effectiveness can be understood across a wide range of possible futures. In the old paradigm, if a decision maker had a problem with an assumption about the scenario used for evaluation of the alternative solutions the entire analysis could be considered invalid. By including "real estate" for scenario variables, the decision maker can ask about any scenario of interest and get immediate feedback on the results.

By including regret in the multivariate scatterplot, an additional degree of depth is obtained. If filtering occurs first at the system-of-systems effectiveness level, the behavior of the remaining systems across many possible futures can be understood. For example, if the effectiveness metric was a success-rate of delivery of supplies through an engagement space, a vehicle like the US Army 2.5 ton truck might show promising effectiveness for scenarios with low enemy activity, but an armored vehicle might show better effectiveness for areas with high enemy activity. Without including scenario variables and regret, it would not be possible to understand where one choice would be better than the other.

4.5.2.1 Limitations

The major caveat for creating the local regret calculation is that the designer must ensure that each possible candidate is tested at each scenario or an optimum performance baseline is generated for each possible future. This is because in the formulation that has been used to this point, each of the thousands of possible design combinations would need to be compared for every point in the possible futures space to obtain a value for local regret. The comparison of such a large number of alternatives would be computationally prohibitive.

If the number of system and subsystem variables is small, implying few actual alternatives under comparison, comparing among the alternatives to calculate local regret is feasible. This is the approach that was taken in the historically-based experiments because only a limited number of platforms existed during the Gulf War. However, in using this technique for design, when thousands of alternatives are being compared, it is probably more feasible to calculate an "optimum" candidate for each possible future and then a simple mathematical difference would create the local regret. After identifying regions of the design space of interest, the regret calculation can be rapidly updated to reflect comparison of each of the candidates for every point in the future scenario space.

4.5.3 Demonstration with Parametric Scenario Generator Data

Another application of the multivariate scatterplot views of the parametric scenario generator is the identification of regions of possible futures with unacceptable system performance. Figure 79 shows the complete, uncolored possible futures space created by the parametric scenario generator. As an example of exploring areas of unacceptable performance, let us consider scenarios which result in a low probability of destroying the target, and a high probability of losing the blue aircraft. For this example, more than an 80% chance (20 losses out of 25) of losing the blue bomber and less than a 50% chance of destroying the target (13 out of 25 unsuccessful bombing runs) will be considered the unacceptable region. This region can be selected by applying filters on the data as shown in Figure 80.

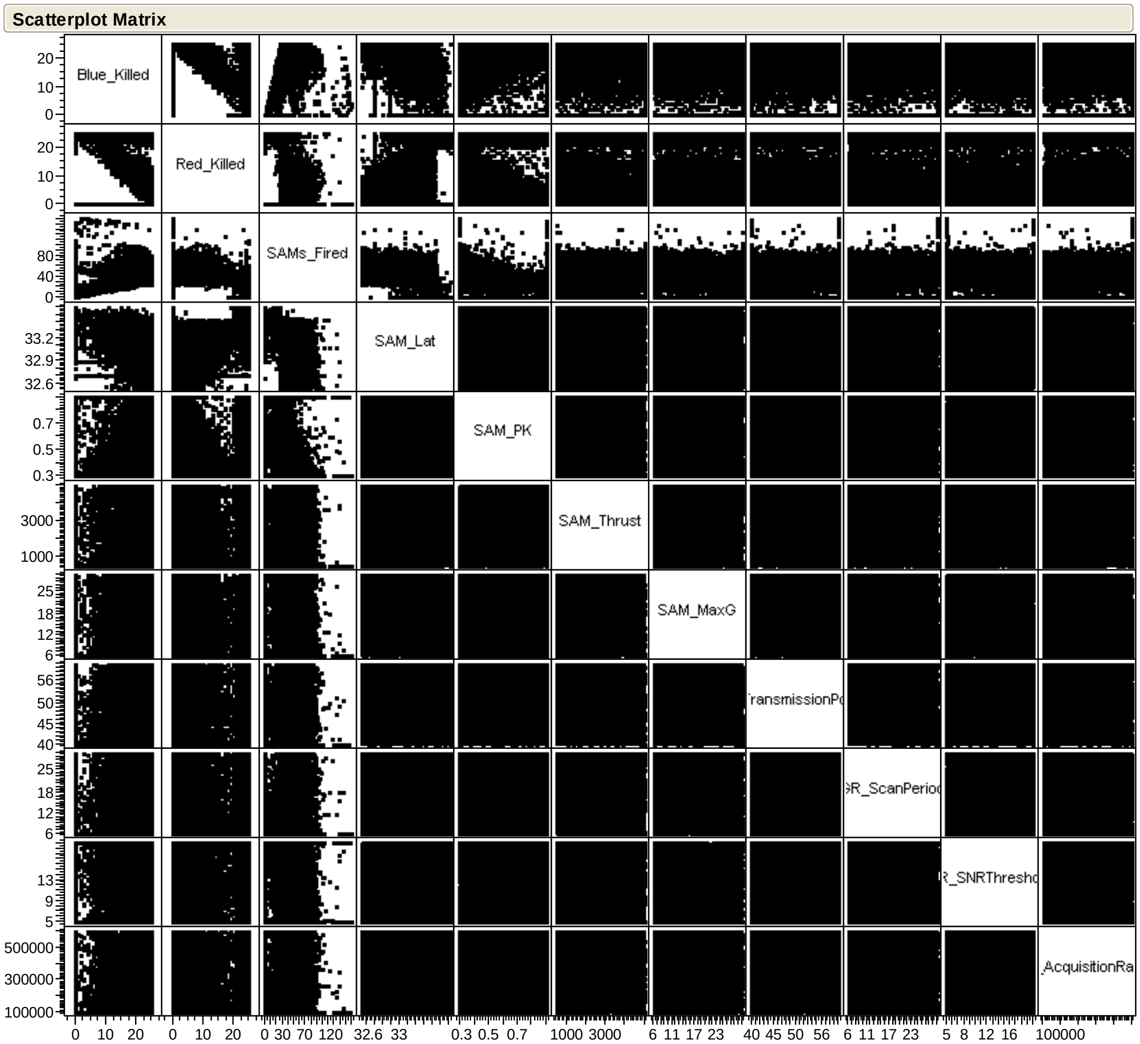


Figure 79: Complete Random Scenario Space

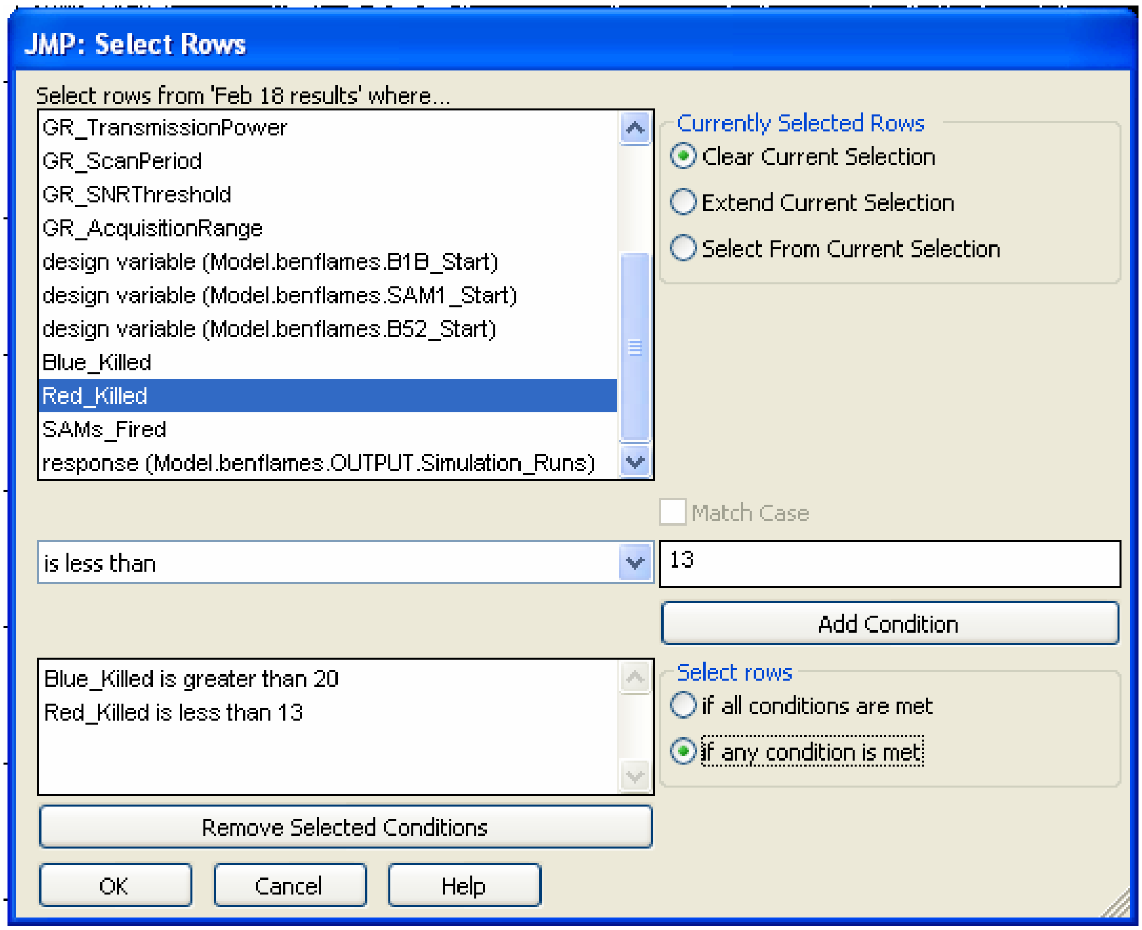


Figure 80: Filtering for a Region of Unacceptable Performance

Once the region of unacceptable performance is selected, all other cases are hidden from the views of the multivariate scatterplot. The remaining points, shown in Figure 81, are the possible futures that result in the region of unacceptable performance. The first result that is immediately apparent from the filtering is the reduction in range of the distance from which the SAM site was avoided. This indicates that if the unacceptable performance only occurs when the blue aircraft comes within the new range of points in the SAM distance row and column. The second result from the filtering has to do with the rest of the points that are left in the other system boxes of the multivariate scatterplot. Because these parameters after filtering still display points over their full range, it means that it is possible to find regions of unacceptable performance for all values of scenario parameters used in the creation.

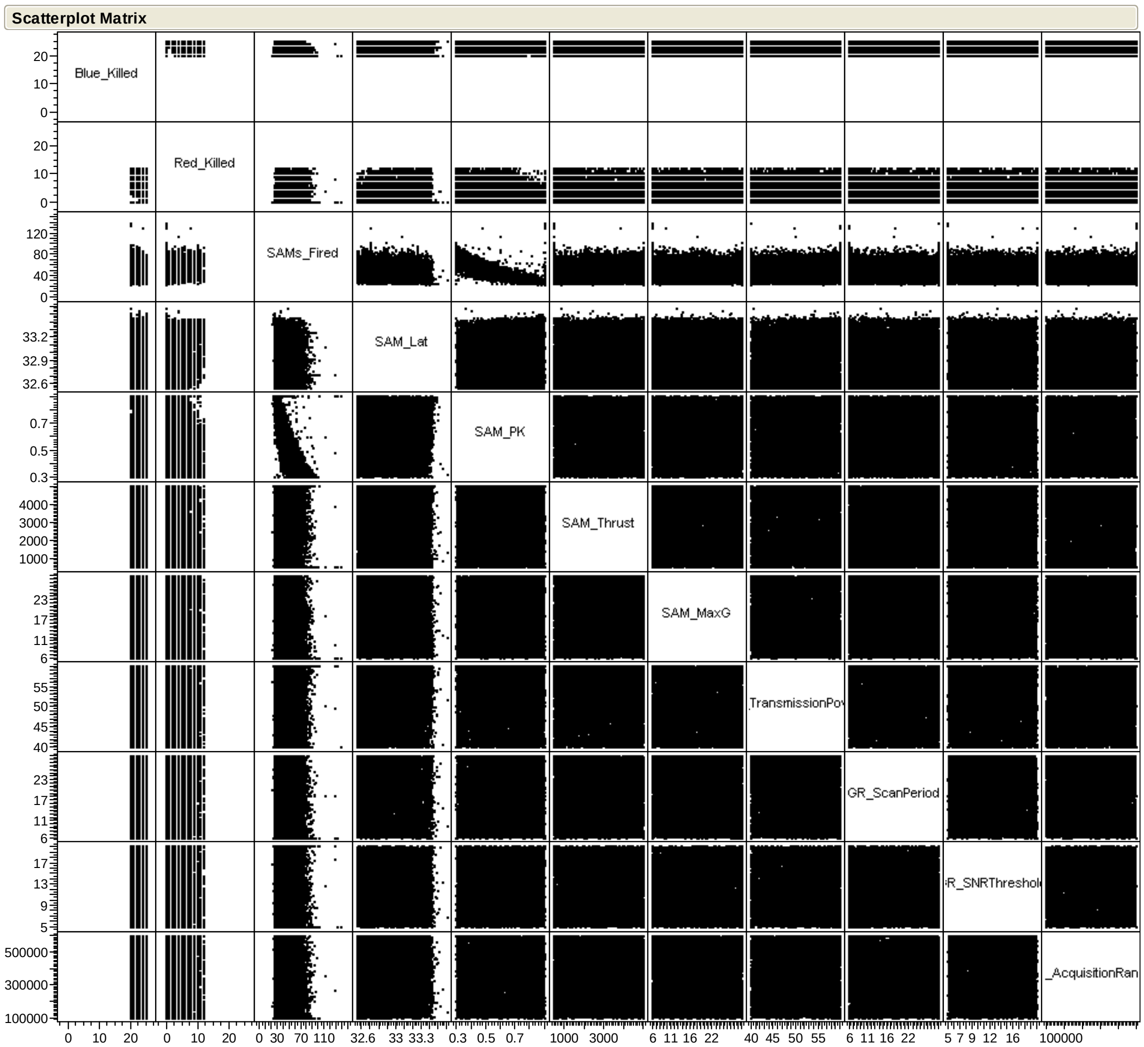


Figure 81: Filtered Results Showing Only Unacceptable Performance

If one additional condition is applied to the multivariate scatterplot, showing only points at a medium-range from the air defense site, a trend appears in the radar systems that can yield unacceptable system performance. Figure 82 shows a much denser clustering of points in the region of high ground radar transmission power and low signal-to-noise ratio threshold than at low ground radar transmission power and high signal-to-noise ratio threshold. This low density of points indicates that, at medium ranges, aircraft were more survivable when the radars with less capability were used by the adversary, regardless of missile performance.

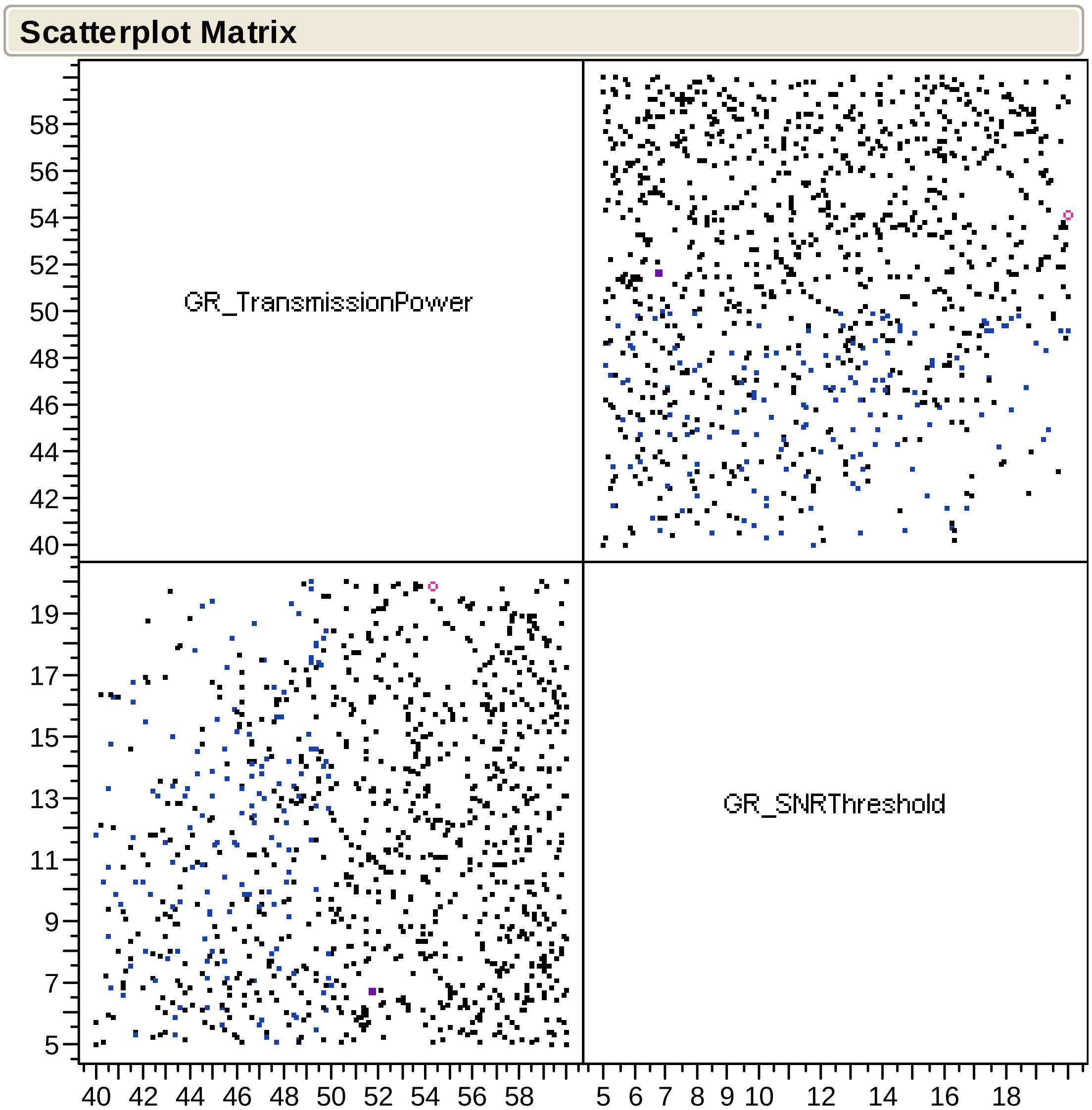


Figure 82: Unacceptable Performance at Medium-Range from SAM Site

If all ranges of aircraft from the SAM site are included in the scatterplot (still under the unacceptable performance filtering) and a filter for relatively low-performance radar is included, the result of the filtering is as shown in Figure 83. This filtering creates trends in the missile performance, as shown in Figure 84. The SAM maximum g-loading still shows points across nearly the entire region of the space, but there is a region of low SAM probability of kill and low SAM thrust that has significantly fewer points than the rest of the plot. The lack of points in the area of low performance indicates that when a low-performance SAM is coupled with a low-performance radar, the blue aircraft is more likely to strike its target and survive the mission.

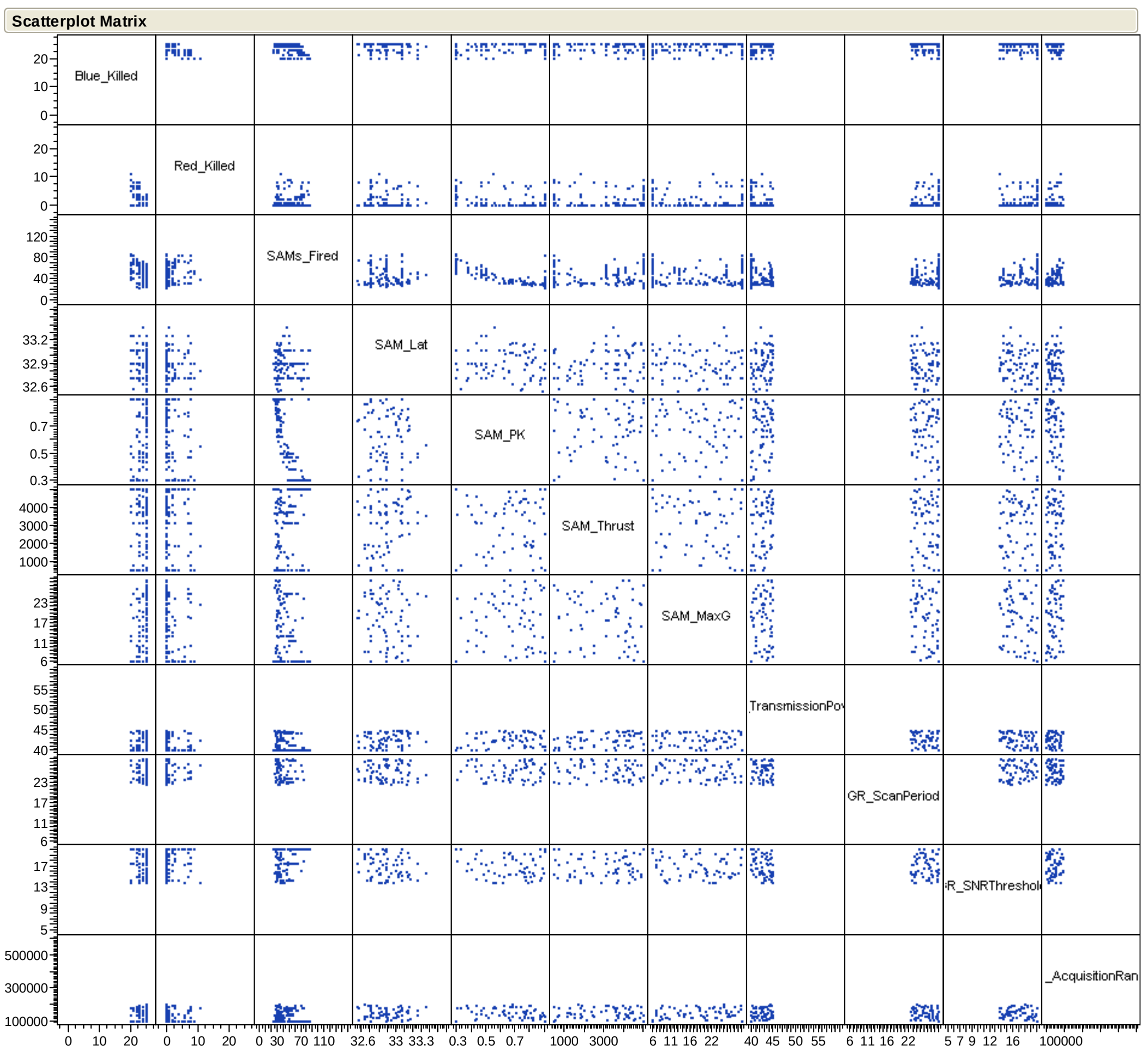


Figure 83: Low-Performance Radar and Unacceptable SoS Performance

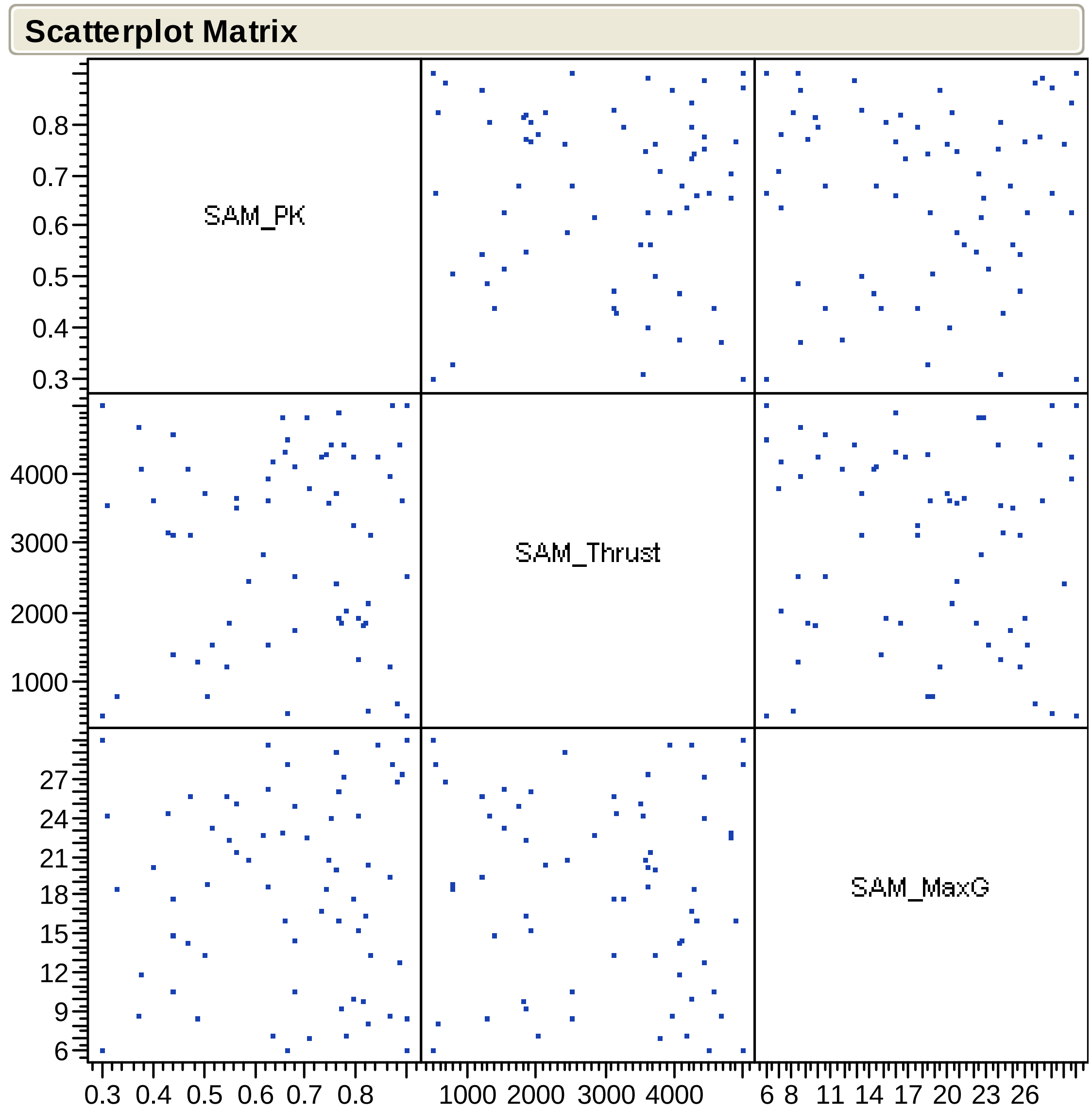


Figure 84: Missile System Parameters for Low-Performance Radar

The identification of these areas can lead to the development of new systems to address shortcomings, or at least keep decision makers informed of situations where results will fall short of expectations. In many cases these results are intuitive, however, intuitive results that reflect the expected performance of real-life systems builds confidence in the ability of the parametric scenario generator to accurately model reality.

4.6 Experimental Summary

The experiments conducted for this dissertation support the hypotheses presented in the previous chapter. FMCS has been shown as an approach that can be used to explore not just a complex design hierarchy space, as Biltgen and Ender showed, but also a scenario space. Using this technique, understanding may be gained as to the performance of various system or system-of-systems alternatives across a wide region of possible futures.

In order to populate the FMCS scatterplots and quantify results, surrogate models were built on data about candidates' performance in possible future scenarios. These surrogate models, which had accuracy of around 90 percent, allowed the population of the multivariate scatterplots rapidly, and in areas of the scenario space where samples were not taken. Additionally, the surrogate models allowed the definition of a continuous function describing the behavior of the candidate's performance over the scenario space. This continuous function was coupled with conventional regret analysis to create Global Regret, a measure of the robustness of the candidates.

To create the data necessary for Global Regret Analysis, a flexible, parametric scenario generator was created. The parametric scenario generator allowed the rapid definition of a scenario of interest. The use of the parametric scenario generator was demonstrated for a single point of interest, a set of randomly generated scenarios, and finally a DoE.

The final experiment mentioned in the experimental setup will be left for the conclusions of this dissertation. It will involve the comparison of the Global Regret Analysis approach to the other robustness methods described in the first chapter.

CHAPTER 5

METHODOLOGY DEMONSTRATION

The proposed method for assessing the robustness of candidates in early design phases has the potential to mesh well with any design methodology that incorporates modeling and simulation that can be run with a DoE. For the demonstration of the Global Regret Analysis Method in the context of a complete alternatives comparison, one particular methodology architecture will be used, but that does not preclude the use of a different methodology approach. The tools used for each step of the general engineering tasks discussed in the research formulation must be selected based on their appropriateness for the particular application; therefore the methodology itself will change based on the problem at hand.

Figure 85 shows a sample matrix of some alternatives for fulfilling each of the general tasks for the Pre-Milestone A phase of the defense acquisition process. Even with this fairly limited set, there are 338,688 possible methodologies that could be constructed, if a single method were chosen from to complete each task. However, in most cases the task will be completed with a variety of methods, or possibly hybrid methods.

Task	Alternative 1	Alternative 2	Alternative 3	Alternative 4	Alternative 5	Alternative 6	Alternative 7
1. Establish the need	User Query	Technologist Query	Gap Analysis				
2. Define the problem	Utility Curves	Functional Flow Block Diagram	Scenario and Environment Definition	Requirements Tracing	Kano Method	Interrelationship Diagraph	Cause and Effect Diagram
3. Establish MoPs and MoEs	Tree Diagram	GOTChA	Affinity Diagram	Pugh Diagram	Prioritization Metrics	QFD	
4. Generate architectures	DoDAF	FEAF	Gartner	MoDAF	TOGAF	Zachman	
5. Generate alternatives	Swarming	Morphology	Brainstorming	BOGSAT			
6. Analyze alternatives	Expert Query	Structure Simulation	Continuous Simulation	Discrete Time Simulation	Discrete Event Simulation	Build & Test	Surrogate Models
7. Compare results	AHP	OEC	TOPSIS	Pareto Frontiers			
8. Make a decision	Electronic Design Review	BOGSAT	Single Point Decision	Report - based			

Figure 85: Subset of Methodology Alternatives

For the Persistent Precision Strike analysis, a methodology was constructed by selecting one, or in some cases two, methods to complete each task. The actual flow of the Georgia Tech Revolutionary Hunter-Killer work is shown in Figure 86; however, because of the sensitive nature of many of the analyses and the need to maintain focus on the Global Regret Analysis demonstration, the methodology in Figure 87 will be presented in this dissertation. The methodology presented herein conveys the important aspects of the Persistent Precision Strike analyses, without details that are irrelevant to the Global Regret Analysis demonstration or were conducted by groups outside the ASDL at Georgia Tech.

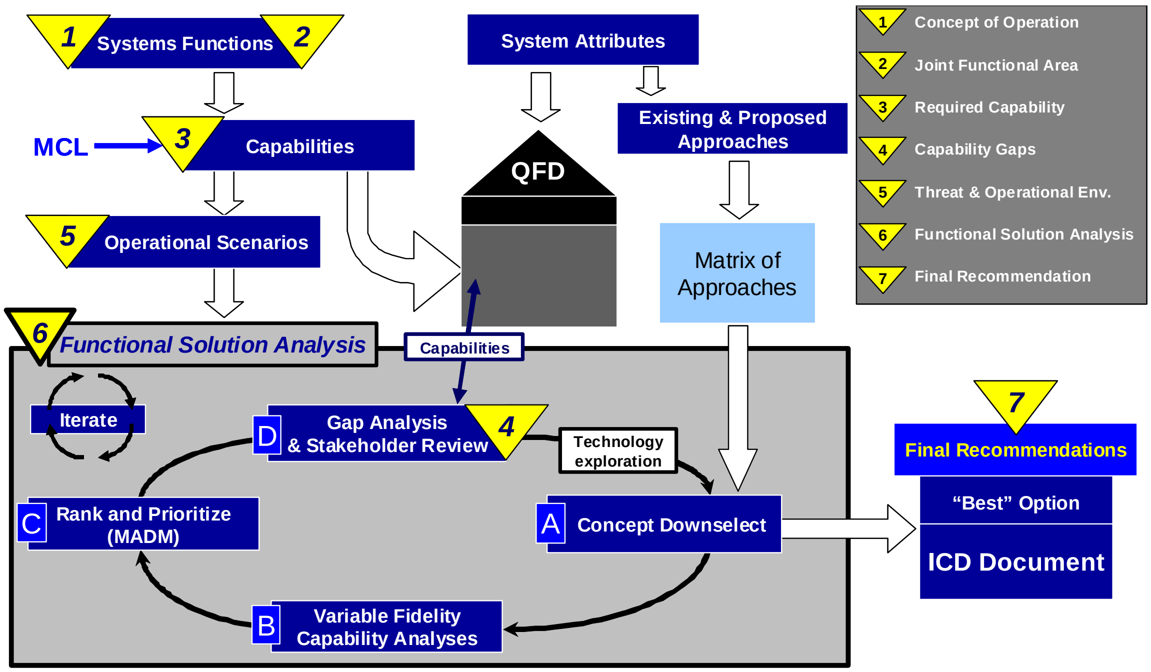


Figure 86: RevHK FSA Approach

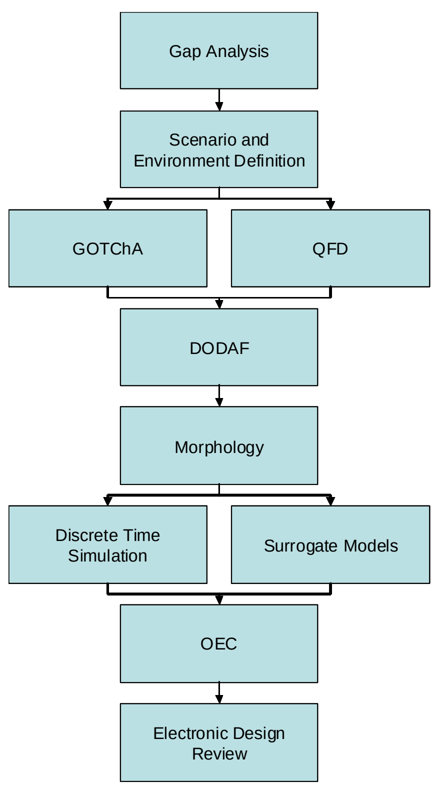


Figure 87: Simplified RevHK Approach

The selection of each of the elements shown in Figure 87 was a joint process between the ASDL at Georgia Tech and the AFRL. Because the ASDL served a support function in the analyses for the Pre-MS A JCIDS process, many of the tasks had already been addressed by the AFRL or other contractors. In these cases, the method for completing the particular Pre-MS A task had already been selected and was beyond the control of the author and researchers at ASDL. This was the case for the Gap Analysis used for completion of Task 1, the QFD used for Task 3, and the DoDAF approach in Task 4. In Task 3, the QFD process undertaken by the AFRL was supplemented with a GOTChA analysis conducted at ASDL. The GOTChA process allowed the understanding of the hierarchy of MOEs and MOPs used in later analyses. The remaining tasks in which the ASDL had influence on the method chosen were tasks 2,5,6,7, and 8.

Task 2 was completed using a collaborative scenario and environment definition. This approach was chosen because AFRL had a specified set of scenarios of interest for the Rev HK. Because of the existing set of scenarios, ASDL used a functional/physical decomposition approach to describe the scenarios and define the problem.

Morphology was chosen for Task 5, in the form of an Interactive, Reconfigurable Matrix of Alternatives, because it possessed several advantages over the other alternatives in the matrix of approaches. While SWARMING, brainstorming, and BOGSAT are all accepted approaches when addressing evolutionary concepts or less complex systems, the complexity of the Persistent Precision Strike SoS necessitated a more systematic method. Because the three methods listed do not necessarily approach the SoS systematically, there is great possibility for potential solutions to be overlooked. However, the decomposition and alternatives enumeration procedure in morphology overcomes the challenges associated with the complexity of the SoS and allow combinatorial solutions that might be otherwise overlooked to be identified.

Task 6, the actual analysis of SoS alternatives, was conducted using a time-stepped simulation and surrogate models. This approach was selected because of the desire to demonstrate the Global Regret Analysis methodology and was conducted independently from the analysis work for the Rev HK research. The discrete time-stepped simulation was chosen for similar reasons to the agent-based simulation in the hypothesis testing chapter of this dissertation, and the surrogate models were used to create the parametric representation of the scenario space. While it would have been possible to use another simulation approach, the surrogate models are a requirement of the Global Regret Analysis methodology.

The use of an OEC for Task 7 was chosen based on the requirements for the Global Regret Analysis methodology. TOPSIS or AHP could have also been imbedded within the Global Regret Analysis methodology, as they both return a fitness value for each concept. However, the differences between these approaches are nuanced and the OEC has widespread acceptance in the defense acquisition community.

Task 8, the actual decision, was hypothetically selected as an electronic design review to showcase the capability of Global Regret Analysis. However, this step was not actually conducted with decision makers because the analysis was conducted in a sanitized manner free from sensitive or dimensional data.

5.1 Selection of an Application

The USAF Persistent Precision Strike was selected for this research based on the interest in the military community in UAVs and Unmanned Combat Aerial Vehicles (UCAVs), the enduring low-intensity conflicts in Iraq and Afghanistan, and the existence of research projects at the ASDL

5.2 Persistent Precision Strike Background

5.2.1 Military Interest

The 1996 Scientific Advisory Board recommended that UAVs and UCAVs be used for missions that are "now, for survivability or other reasons, difficult for manned aircraft." [61] All four branches of the US Military have incorporated unmanned systems into their near and far-term plans for military operations. The US Army incorporates a number of sensor and weaponized UAVs and Unmanned Ground Vehicles (UGV) in their Future Combat System. The USMC and USAF both have strategic visions that call for the use of weaponized UAVs and sensor UAVs. In the case of the Marine Corps, these vehicles are indigenous at the platoon or squad level, with weaponized versions being available at the battalion level. [89], [237] The US Navy's Sea Power, Sea Base, and Sea Shield visions all call for remotely operated and autonomous aerial and undersea vehicles.

5.2.2 Brief UAV History

Unmanned Aerial Vehicles (UAVs) have been used by the US military since the USAF deployed the "Lightening Bug" during the Vietnam War. [237] However, it was not until 2002 that armed versions of these remotely piloted craft were employed in US operations. On February 7th of that year, a MQ-9A Predator fired hellfire missiles into a convoy of Sport Utility Vehicles (SUVs) belonging to the Al-Qaeda terrorist network, killing a senior Al-Qaeda member. This event represented the first time in history that an air strike was carried out by a remotely piloted surveillance vehicle, with almost no risk to members of the US Military.

In the past 5 years since the 2002 Predator air strike, the state of the art in UAVs has advanced dramatically. An armed successor to the Predator, the MQ-9B Reaper has the ability to carry significantly more payload and stay aloft longer. The Israeli Harpy system can be launched from a truck and will autonomously hunt and kill air defense systems. The Global Hawk reconnaissance platform is truly remarkable; with an endurance of over 35 hours, the ability to fly unrefueled half way around the world, under either remote or autonomous control [238], this platform effectively replaces the dangerous U-2 capability of the Cold War. According to Larry Dickerson, in many countries the UAV is viewed as a cheaper alternative to satellite surveillance systems, with the ability to duplicate many of their capabilities. [93]

The advantages of using unmanned systems in military operations are numerous. The following table (Table 17) summarized the advantages and disadvantages of UAVs outlined in the US Air Force's Strategic Vision for Remotely Piloted Vehicles (RPVs) and UAVs as well as those from other sources. [237]

Table 17: UAV Advantages and Disadvantages

Advantages	Disadvantages
Operation for longer than human endurance allows	Integration into existing airspace difficult
Operation in chemically, biologically, or radioactively contaminated environments	Data bandwidth limitations
Reduced ground crew operational tempo because of endurance	Weight limitations similar to manned systems
Reduced wear and tear because of fewer takeoff/landing cycles per flight hour	Weather limitations
Crews do not necessarily have to deploy forward to operate vehicles	Reliability issues
Reduced operational logistics, support, and cost footprints	Susceptibility to Jamming
Off-loading mundane tasks through machine autonomy	Organizational issues, acceptance in "pilot-centric" culture
Expansion of traditional flight and altitude envelope	Data fusion abilities (Elliot)
Reduce risk to pilots (Buxbaum)	

5.2.3 Historical Persistent Strike

5.2.3.1 Tacit-Rainbow

Tacit-Rainbow was a program during the 1980s designed at developing a Persistent Anti-Radiation Missile (PARM) for the SEAD mission. The goal of the Tacit-Rainbow program was to create a platform that would be launched in large numbers in advance of a bombing raid. The platforms would loiter in the area, and autonomously attack any radiation emitting devices within a certain frequency band (radar sites). The primary advantage of the loitering PARM was that if a radar site attempted to protect itself by shutting down, the PARM would simply wait for it to turn back on again. [171]

Tacit-Rainbow was unique in that once, launched it did not need targeting instructions from the airman who launched it. Rather it would autonomously loiter in a preprogrammed area and wait for a target that met a certain set of criteria, preprogrammed into the system.

5.3.2.2 Harpy

The Israeli Israel Aircraft Industries Harpy system is similar in purpose to the Tacit-Rainbow. The Harpy is launched from a modified truck chassis, moves to a loiter area, and begins searching for radar emitters. If a radar emitter is detected, and determined to be a target by the Harpy's logic system, the aircraft will attack the emitter. While not officially designated as such, this fits the definition of PARM. The Harpy is a "fire-and-forget" weapons system [124], which means that once it has been launched, it will behave autonomously until it either runs out of fuel, or destroys itself attacking a target.

Both Tacit-Rainbow and the Harpy are autonomous systems that use preprogrammed logic to decide whether to attack a detected target. However, the level of intelligence required for these vehicles is limited because of their high level of specialization and the relatively unique nature of their targets. Radar systems are used for a wide variety of applications from range finding and speed detection to aircraft tracking. However, the bandwidth and power associated with military radar systems makes them fairly unique. Only commercial aircraft control systems

5.2.4 Current USAF Fleet

Very few US military systems have the ability to maintain station over an area, find a target in that area, and then deliver a weapon to the target. Killbox Interdiction techniques [138] achieve a similar capability by employing multiple aircraft or combinations of aircraft and ground systems. In a killbox mission, strike aircraft loiter in an area of interest, and are assigned targets by an Intelligence, Surveillance, and Reconnaissance (ISR) aircraft, or possibly by troops on the ground. This approach is reminiscent of using infantry to scout for artillery, then having the infantry radio coordinates and adjustments to the battery. However, the separation of the ISR and strike responsibilities increases the exposure of units and coordination required.

The USAF currently operates the following manned ISR platforms.

- E-3C Airborne Warning and Control System (AWACS)
- E-8C Joint Surveillance Target Attack Radar System (J-STARS)
- EP-3 (Aries II) Navy
- RC-12 (Guardrail) Army
- RC-135 (Rivet Joint)
- U-2 (Dragon Lady)

Additionally, the USAF operates three unmanned ISR platforms, the Predator, Reaper (Figure 88), and Global Hawk (Figure 89 US Museum of the Air Force). The operating costs of these aircraft vary widely relative to the manned ISR assets. A 2005 assessment by the Naval Research Advisory Committee also found that operating costs for the Global Hawk were higher than for any existing or proposed Navy surveillance aircraft. The advisory committee, a group of independent civilian scientists who advise the Office of Naval Research, determined the cost of operating the Global Hawk at $26,500 per flight hour. The group also reported operating costs for the Predator at $5,000 per flight hour. In comparison, the group set the Navy's cost for operating its E-2C Hawkeye, a manned airborne warning and control aircraft, at $18,700 per flight hour. [22]

Figure 88: MQ-9B Reaper

Figure 89: Global Hawk [238]

5.2.4.1 Strike

The USAF currently has 8 strike aircraft available in its inventory (Figure 90). [241] Because of the precision nature of the persistent strike mission, the payload capacity of the strike platform is not considered a driving factor. However, the endurance and operation cost for aircraft that will spend the majority of their time loitering while waiting for target information is important for persistent strike effectiveness and cost.

Figure 90: USAF Attack Aircraft

Endurance is a critical constraint on the strike component of the killbox interdiction mission. While aerial refueling can extend the endurance of a loitering strike platform, it does not address all problems with manned strike systems. First, when refueling the strike platform must either depart the killbox, or at the very least leave station to rendezvous with the tanker aircraft. This creates gaps in the strike coverage that must be filled by another aircraft or considered mission downtime. Secondly, even with refueling the manned system is limited by pilot endurance. While concrete numbers on pilot endurance are hard to find, a typical value for a single seat fighter is probably around 8 hours.

The cost per hour of loitering a strike aircraft is significant. The F-16 is favored by the Air National Guard because it has a relatively low operating cost of somewhere around $4000 per hour. [98] Platforms such as the B-52 have operating costs (largely driven by maintenance) which are significantly higher. The endurance limits, high cost, and dull-nature of the killbox mission make UAVs a good candidate for filling the role of manned strike platforms.

The USAF has a wide variety of ground attack munitions at its disposal for use on an unmanned aircraft as shown in Figure 91. Additionally, the US Army has demonstrated that its Hellfire missile can be successfully employed on UAVs. The Predator B currently employs Hellfire missiles.

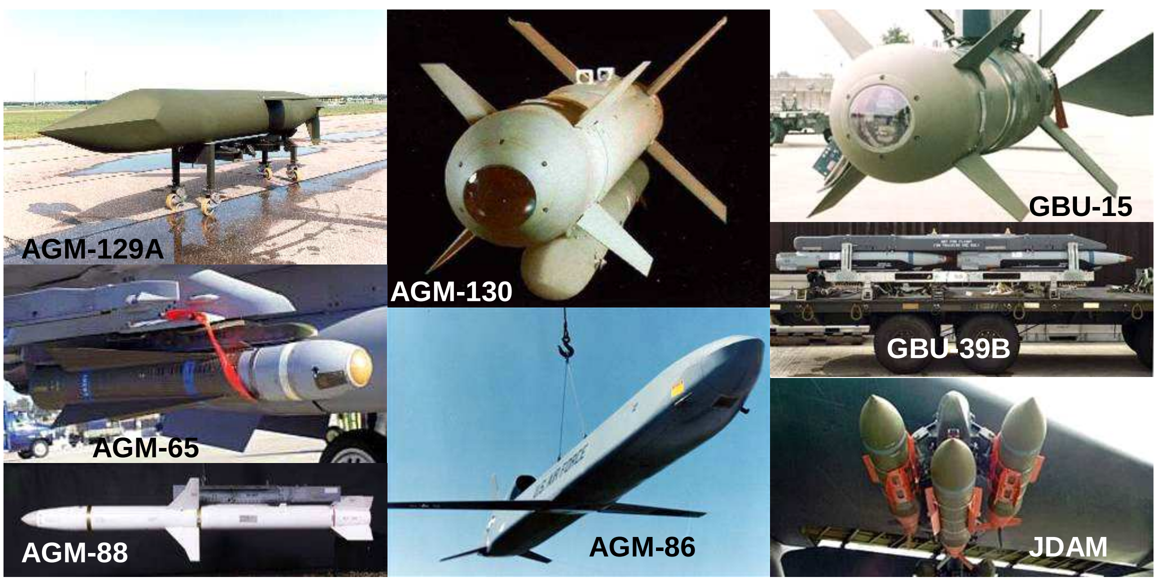


Figure 91: USAF Air to Ground Attack Munitions

5.3 Approach

5.3.1 M&S Environment

A number of Modeling and Simulation (M&S) environments were considered for use in the analysis of candidate Revolutionary Hunter-Killer alternatives. Because the problem under consideration is similar to the assessment of candidates in the hypothesis testing section of this dissertation, a time-stepped environment will be sought for the same reasons outlined for the other assessment. The following modeling environments were considered.

- FLAMES by Ternion
- SEAS by DoD
- NetLogo by Center for Connected Learning at Northwester University
- ATMAS by Diana Talley at Georgia Tech's ASDL
- MATLAB by MathWorks

5.3.1.1 Selection Criteria

Each of the five M&S environments listed above have benefits and drawbacks. In order to objectively compare them, the following criteria were used. They are discussed in order of importance, with the first criterion being most important.

5.3.1.1.1 Availability of Code

The most important criterion for selecting an M&S environment was the availability of the code to the researcher. Many commercial M&S environments have high costs associated with licenses, so it was important to ensure that the code was free, low-cost, or already licensed for use by the researcher. In the case of each of the M&S environments listed, it was possible to use the codes in the laboratory setting, though in the case of FLAMES, required using a specialized terminal for model development. MATLAB, ATMAS, and SEAS all existed in the laboratory environment and could be loaded onto a personal terminal. Finally, NetLogo was available as freeware and could be downloaded to laboratory or personal computers.

Because its freeware status, NetLogo was considered the best in terms of availability; it was followed by MATLAB, ATMAS, and SEAS, which were considered equally desirable. The FLAMES package was considered least desirable because of the requirement to use the special terminal for model development.

5.3.1.1.2 Suitability for Modeling SoS

The investigation of the concepts for the Revolutionary Hunter-Killer will include interactions between systems working together to achieve a capability. In order to successfully assess the robustness of the various candidates, the M&S environment must possess the ability to model those interactions. All of the M&S environments have some ability to model the interactions of SoSs. ATMAS was built in, and executes in, MATLAB, demonstrating this capability. However, because of the procedural nature of the MATLAB programming language, dealing with SoS is more difficult that with a more object-oriented approach. FLAMES, SEAS and NetLogo all use a more object-oriented approach than MATLAB, and were considered more desirable for that reason.

5.3.1.1.3 Existing Knowledge or Shallow Learning Curve

The desire to complete the modeling tasks relatively quickly drove the search for a code that either had a wide base of existing knowledge that could be leveraged, or a relatively shallow learning curve. In the engineering community, MATLAB is perhaps one of the best well known programming languages and environments. The author has used MATLAB extensively in the past, and therefore it was a very desirable choice from the standpoint of existing knowledge. FLAMES, SEAS, and ATMAS each have been used in research projects with which the author was affiliated, though not extensively by the author. The existence of the knowledge within the laboratory community was a positive mark for each of these environments.

Unfortunately, the FLAMES package has a steep learning curve which requires significant time before productive models can be created. The work earlier in this dissertation leveraged existing models in many places, but that was not possible for the Revolutionary Hunter-Killer study. The steep learning curve associated with the FLAMES package detracted from its attractiveness as an M&S environment for the Revolutionary Hunter-Killer study. SEAS and ATMAS both were unfamiliar to the author, but possessed moderate learning curves, especially relative to FLAMES. This made them more attractive than FLAMES, but less so than MATLAB. Finally, NetLogo possessed an extremely shallow learning curve, a few days to develop fairly advanced simulations. The extreme simplicity of the language made NetLogo the most desirable M&S environment other than MATLAB.

5.3.1.2 Selection

Based on the evaluation criteria discussed above, the author selected NetLogo as the best compromise solution for the M&S environment. While the coding language was not already known, the appropriateness of the environment for SoS problems, the ability to install the program on any computer and the shallow learning curve made it the best choice overall.

5.4 Application of Methodology

Methodology application required working through each of the eight general tasks for the Pre-Milestone A phase of the defense acquisition process. In some cases, because the task would be completed by parties other than the analyst, existing results from the public domain were used.

5.4.1 Establish the need

The US Air Force has already established persistent, precision strike (engagement) as one of their priorities in the future vision (Figure). [24] During Operation Allied Force, the minimum time it took to coordinate high altitude ISR assets with a strike platform was 12 minutes [226]. However, the USAF has a goal of a single-digit minutes for the kill chain. [116]. This goal provides the general framework for the development of the Revolutionary Hunter-Killer System.

Figure 92: Air Force Vision [24]

The body of work that truly established the need for the Revolutionary Hunter-Killer included two gap analyses, and was conducted within the USAF. [198], [199] According to Bowman, the capabilities are being pursued through the program are the surveillance of an area-of-interest for time-sensitive targets, and the prosecution of those targets. [30] These capabilities were used as a baseline for the Next Generation Morphing Aircraft Structures program, which combined with the gap analyses were the predecessors of the Revolutionary Hunter-Killer. [30], [31]

Because the need for the Revolutionary Hunter-Killer had been established by the Air Force prior to the start of this work, additional justification for the need will not be pursued.

5.4.2 Define the problem

The analysis of Revolutionary Hunter-Killer alternatives should allow the researchers to guide further research with a more clear understanding of trades at the SoS level. Because the Revolutionary Hunter-Killer is still in the pre-conceptual design phases, before the solidification of the Initial Capabilities Document, the SoS level trades could also be described as an Analysis of Approaches. The approaches mean high level trades among large classes of systems, as opposed to limited system level trades.

Of particular interest for this work is understanding the impact of a single, very capable (and presumably expensive) vehicle, versus a team of moderately capable vehicles, versus a swarm of low-capability vehicles. Understanding the capability of a single vehicle falls very much in line with traditional vehicle analysis, but the team and swarm concepts both rely on emergent behavior of the group for capability. This emergent behavior is not immediately apparent from the specifications of the individual vehicles, which is often the focus of the AoA in the post-ICD analysis. Therefore, understanding these trades upfront in the pre-ICD phase is important to guide the more vehicle-centric AoA.

5.4.3 Establish MOPs and MOEs

The MOEs for this problem should quantify how well the different alternatives are able to complete the mission and how much is costs to complete. At the highest level, this can be described for the kill-box mission in the following questions.

- What percentage of the available targets did the SoS find and kill?
- How quickly were the available targets found and killed?
- How many times did the aircraft have to resupply in the process?
- What was the cost of operating those systems?

The MOPs at this stage of analysis are not as important as the overall capabilities of the aircraft. However, two items are of interest and fall naturally out of the capability considerations: fuel capacity and weapons capacity.

Because an excess of either fuel or weapons will have an adverse effect on vehicle weight and, consequently, cost it is important that the aircraft have an appropriate amount of fuel and number of weapons. The appropriateness of the fuel and weapons payloads will be monitored by recording the reason for each resupply. By using this metric of performance, the driving factor can be identified whether it is fuel or weapons.

5.4.4 Generate architectures

Because the Revolutionary Hunter-Killer will be designed to fill a kill-box interdiction role, the basic architecture of the current unmanned vehicle that fills this role, the Reaper, will be used. This basic architecture includes a vehicle or vehicles, under the control of a ground station, working to find and then deliver a missile against a target. A representation of the architecture and Concept of Operations (CONOPS) is shown in Figure 93. This basic structure was identified based on Bowman's discussion of kill-box interdiction.

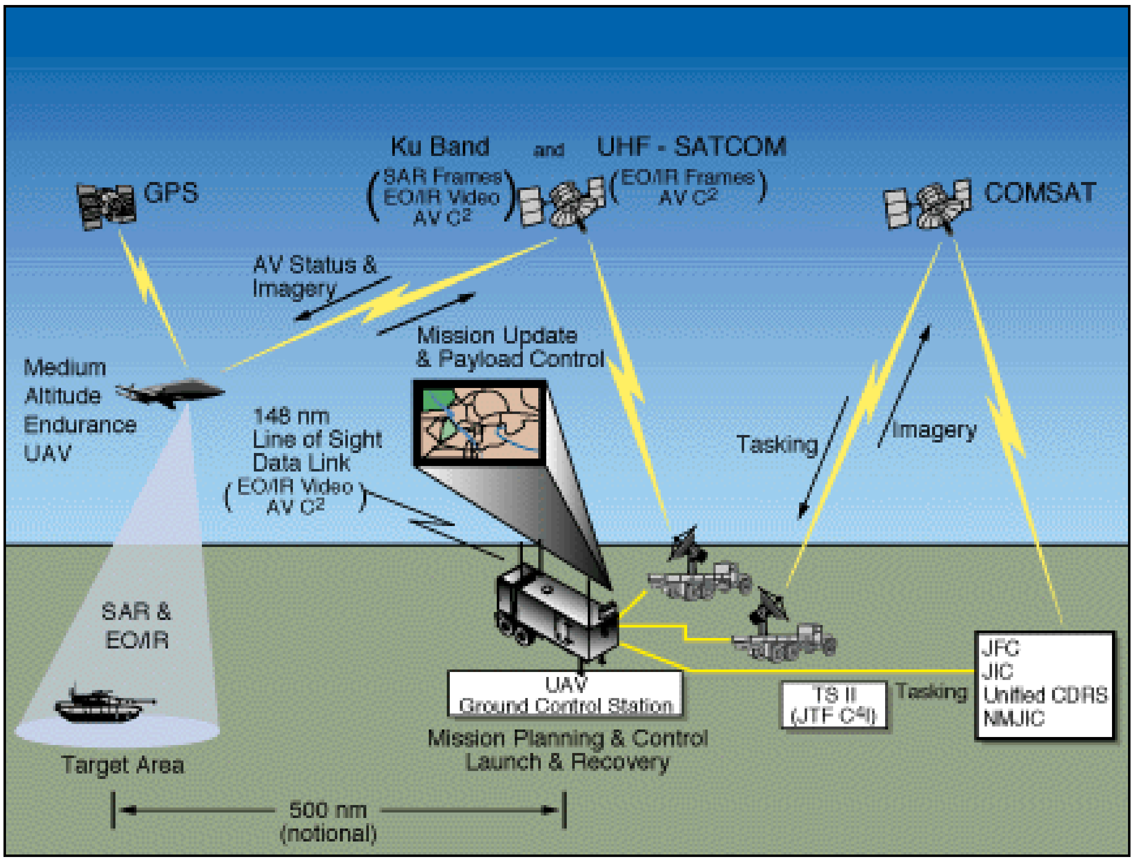


Figure 93: CONOPS [81]

The level of autonomy assigned to the unmanned vehicles in the kill-box interdiction mission has the potential to impact the architecture of the SoS. If vehicles are allowed to operate completely independently, there is no need for a remote pilot ground station, wide bandwidth communication or pilot training. However, this level of autonomy increases the requirement for on-board computation and machine intelligence. At the other end of the spectrum, a completely remotely piloted aircraft requires a large amount of bandwidth for pilot awareness, extensive ground control stations, and provisions for handling lost communications with the aircraft.

Because the architecture is somewhat dependent on choices for the alternatives, the basic architecture will be assumed, but only the elements operating in the kill-box will be modeled. Other elements, such as aerial refueling and ground station response times, will be modeled by through parametric times for communication responses and re-supply. This assumption allows the impact of various architectures to be considered on the in-kill-box capability, but still allows the analysis to remain in the scope of this dissertation.

5.4.5 Generate alternatives

5.4.5.1 Approach

Morphological analysis was selected as the method for identifying the possible alternatives for fulfilling the kill-box interdiction mission in a systematic way. These methods have gained popularity in the aerospace industry in recent years as a way to deal with the massive size of the possible design space. Morphological analysis provides a "method for identifying and investigating the total set of possible relationships or 'configurations' contained in a given problem complex." [193] Because this method was developed in the middle part of the 20th century, the computational resources were not available to provide significant numerical analysis for problems of the scale seen in conceptual design space. Therefore the focus is to reduce the initial set of alternatives to a manageable set by applying filters to the conceptual design space.

The morphological analysis creates an n-dimensional matrix where each dimension corresponds to a particular physical or functional feature of the system or system of systems. In this construct, each member of the space of alternatives would correspond to a cell in the hypercube defined by the n-dimensional matrix. Morphological analysis then removes incompatible combinations in this matrix and then applies constraints to the various dimensions of the problem in an effort to reduce the number of alternatives to a manageable set that can be evaluated with quantitative methods. In recent years, effort has been made to incorporate limited quantitative analysis in the framework of a morphological analysis. [80]

In any conceptual design problem, there exits the possibility of trillions of design alternatives that have the potential to satisfy the requirements of the problem to varying degrees. In order to systematically assess these alternatives, the ASDL at Georgia Tech has created a tool called the Interactive, Reconfigurable Matrix of Alternatives (IRMA). The IRMA allows experts from various disciplines and system designers to evaluate the design alternatives by filtering concepts based on Technology Readiness Level (TRL), cost, performance, etc. These filters, coupled with expert engineering judgment allow the reduction of the design space from trillions of alternatives to a manageable subset that can be further evaluated for concept selection.

The rows of an IRMA represent a physical or functional breakdown of the system of interest, depending on knowledge of the system architecture. In each column, alternatives are listed that could satisfy the functional or physical characteristic of the row. The alternative space is then defined as all feasible combinations of systems which are created by selecting an item from each row. This combinatorial space is the set of alternatives that must be evaluated in the AoA.

Once populated based on background research and expert opinion, a typical Matrix of Alternatives will represent well over a trillion combinations of concepts which must be systematically evaluated to find the best overall concept. (To put this in perspective, one trillion cases evaluated at one case per second would take 31,710 years to evaluate.) One important characteristic of the IRMA is the identification of incompatibilities between characteristic alternatives prior to the evaluation of concepts. Additionally, a Technology Readiness Level (TRL) is assigned to each alternative in the matrix, and concepts can be filtered such that only concepts meeting the minimum TRL will be displayed. This ensures that if a particular option is selected, all other incompatible alternatives are eliminated from the matrix.

In order to begin reducing the design space to a manageable set that can be evaluated using reasonable computational effort, a collaborative meeting is held to begin reducing the options in the IRMA. Initially, a minimum TRL for the project is established based on the resources available and the desired date of system deployment, which then filters alternatives in the matrix that do not meet the minimum TRL. Then the customer and engineers assess each row of the matrix and identify concept alternatives that should be eliminated from consideration due to a major defect with respect to an important measure of goodness. As each row is evaluated, the IRMA updates to reflect the remaining compatible combinations that must be considered. The set of compatible alternatives that remain after filtering are those that must be compared to identify the best system concept.

5.4.5.2 Revolutionary Hunter-Killer Alternative Exploration

The IRMA used for identifying Revolutionary Hunter-Killer alternatives is shown in Figure 94. The first column of the matrix shows the categories for which alternatives were defined, organized by the SoS hierarchy. At the top level are SoS variables, primarily concerned with the mix of hunter, killer, and hunter-killer. The aircraft level contains mission profile characteristics for the aircraft, as well as the presence of survivability enhancing characteristics. The weapon level contains the number and type of weapons that the hunter-killer or killer vehicle carries. Options for the primary and secondary sensor systems and the communication equipment are included in the sensor level category of the matrix.

SOS Level	Selection	OPTIONS						
Configuration Options	H&K	H&K	H/K	H & K & H/K				
Number of H/K	1	0	1	4	10	25	50	100
Number of H	0	0	1	4	10	25	50	100
Number of K	0	0	1	4	10	25	50	100
Aerial Refueling	No	No	Yes					
System Level								
Number of Operators	1	1	2	3	4			
Modularity	Low	Low	Medium	High				
Autonomy	Remote	Manned	Remote	Autonomous				
Aircraft Level (Common)								
Fly (Hunter Mode)								
Operating Altitude (ft)	40000	30000	40000	50000	60000	80000		
Range (nm)	1000	500	1000	1500	2000	2500		
Cruise Speed (Mach #)	0.6	0.2	0.4	0.6	0.8	0.9		
Dash Speed (Mach #)	1.4	0.8	1	1.4	2	2.5		
Endurance (hr)	10	5	10	20	30	40		
Fly (Killer Mode)								
Operating Altitude (ft)	20000	10000	20000	30000	40000	50000		
Range (nm)	50	50	100	200	300	400	500	
Cruise Speed (Mach #)	300	125	200	250	300	350		
Dash Speed (Mach #)	350	125	200	250	300	350		
Endurance (hr)	4	2	4	6	8	10		
Survive								
Avoid Detection	Some stealth tech	No stealth tech	Some stealth tech	Lots of stealth tech				
Avoid Hit (counter measures)	No	No	Yes					
Survive Hit (redundant systems)	No	No	Yes					
Weapon Level								
Primary Weapon								
Weapon 1 Payload (Internal)	External	Internal	External					
Weapon 1 Type	SDB I	SDB I	SDB II	Subsonic Missile	High Speed Weapon	Hypersonic	GBU-28	None
Number of Weapons	1	1	2	4	6	8		
Secondary Weapon								
Weapon 2 Payload (Internal)	Internal	Internal	External					
Weapon 2 Type	Subsonic Missile	SDB I	SDB II	Subsonic Missile	High Speed Weapon	Hypersonic	GBU-28	None
Number of Weapons	1	1	2	4	6	8		
Sensor Level								
Primary Sensor								
Sensor Type	RF	EO/IR	RF					
Sensor Payload Location	Internal	Internal	External					
Number of Sensors	1	0	1	2	3	4		
Coverage	Med	Very Low	Low	Med	High	Very High		
Range (nm)	10	1	10	25	50	100		
Resolution (m)	1	0.1	0.5	1	2	3		
Secondary Sensor								
Sensor Type	EO/IR	EO/IR	RF					
Sensor Payload Location	Internal	Internal	External					
Number of Sensors	1	0	1	2	3	4		
Coverage	Med	Very Low	Low	Med	High	Very High		
Range (nm)	10	1	10	25	50	100		
Resolution (m)	1	0.1	0.5	1	2	3		
COMM								
COMM Type	LoS	LoS	SatCom					
COMM Payload (Internal)	Internal	Internal	External					
COMM Range (nm) [Derived]	100	50	100	150	200	250		
COMM Frequency (MHz)[Derived]	1000	500	1000	1500	2000	2500		

Figure 94: Hunter-Killer IRMA

The rows of the IRMA for the Revolutionary Hunter-Killer were populated based on a literature search conducted by the research team. The literature search included current and proposed systems, and was directed toward creating realistic bounds for the elements of the IRMA. Because the effort was primarily aimed at understanding the difference between a single hunter-killer aircraft and a system of separate hunters and killers, details of the aircraft design beyond the mission profile were not considered. Sensor technology was considered to be very important, as was using an off-the-shelf missile technology.

5.4.5.3 Identification of an Analysis Sub-set

Because of the scope of this dissertation, identification of a small subset of potential alternatives for analysis was important. Naturally, a representative single-aircraft revolutionary hunter-killer should be compared to a two-vehicle system, as understanding the difference is one of the goals of the research. Additionally, understanding how the number of each type of aircraft impacts capability is desirable so that the relative costs can be understood. These high-level fleet sizing and aircraft type questions drove the selection of each of the candidate SoS shown in Table 18. These were selected so as to represent very different approaches and provide insight for recommendations for the initial capabilities study.

Table 18: Revolutionary Hunter-Killer SoS Options

Alternative	**1**	**2**	**3**	**4**	**5**
Brief Description	Highly-capable single HK	Moderate number of hunters, one highly capable killer	Moderate number of less capable HKs	Swarm of hunters and killers	HKs Augmented with Hunters
Number of HKs	1	NA	5	NA	2
Number of Hunters	NA	10	NA	15	4
Number of Killers	NA	1	NA	8	NA
HK Fuel	6785	NA	5999	NA	6399
Hunter Fuel	NA	5450	NA	3917	5580
Killer Fuel	NA	C-130	NA	3532	NA
HK Weapons	4	NA	2	NA	4
Killer Weapons	NA	100	NA	2	NA
Sensor Size	20x20	7x7	10x10	5x5	10x10

5.4.6 Analyze alternatives

Two primary steps were used for the analysis of the five hunter-killer concepts. The first step involved the sizing and synthesis of the particular aircraft that make up the SoS. This step included a cost calculation in addition to energy-based and empirical sizing equations. The "Aircraft Sizing and Synthesis Module" contains the calculations for this step and was created in Microsoft Excel ® with additional Visual Basic scripts run to converge the designs. The second step was the evaluation of the sized alternatives in the parametric scenario generator. Two modules were combined to complete the second step. The MATLAB based "Terrain and Urban LOS Module" calculates the ability of the aircraft to see an area of interest based on urban building and street layout and the mountainousness of the area. The output of this module was fed into the "Mission Analysis Module," where the effectiveness of the particular concept was evaluated. This module was built in the NetLogo environment. The flow of information among the modules is shown in Figure 95.

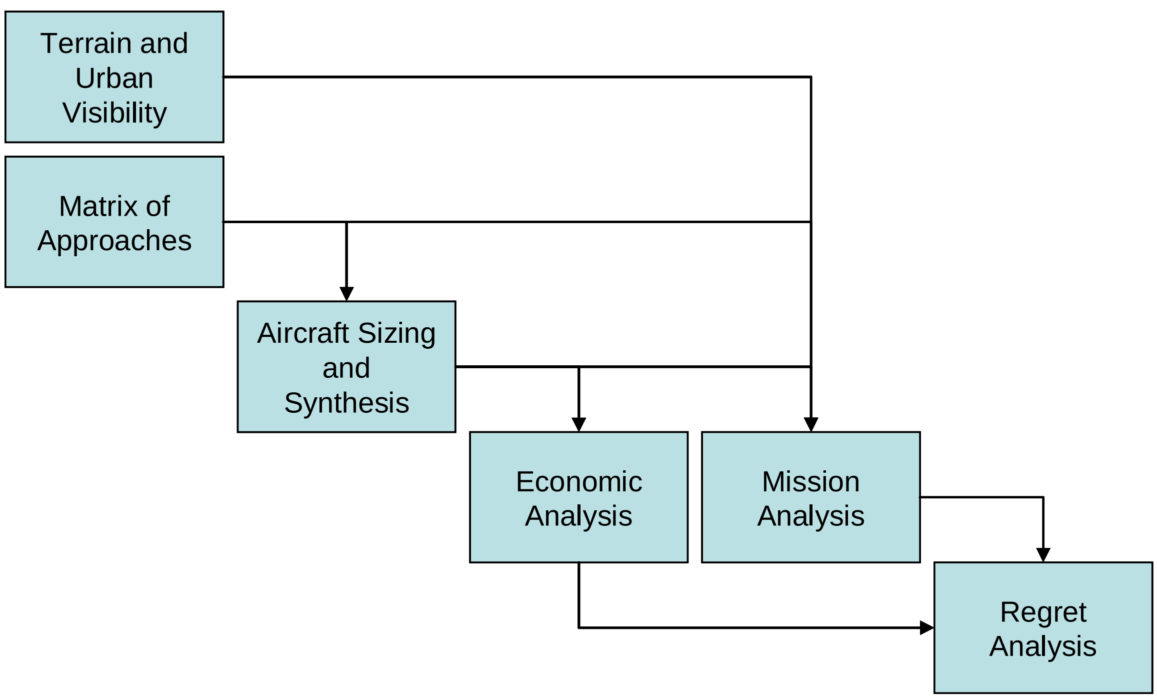


Figure 95: Information Flow Among Modules

5.4.6.1 Aircraft Sizing and Synthesis Module

The aircraft sizing and synthesis module is currently a combination of energy-based constraint analysis, historical engine performance, and historical weight estimation based on a ratio of fuel volume to empty weight. There are seven sheet colored in yellow. Each is described below.

Aircraft Sheet – There are three primary input areas on this sheet, corresponding to aircraft mission parameters, environmental parameters, and aircraft design parameters. There are three primary output areas of the sheet that display the graphical mission profile, the aircraft weight data for the current input settings, and finally a comparison plot of the thrust-to-weight and wing loading of the aircraft relative to three other data points: the F-35 Lightning II, the Reaper UAS, and the Global Hawk.

Aircraft mission parameters are linked from the main MOA page, but can be manually adjusted on the Aircraft sheet. The parameters associated with the attack mission segment are not part of the MOA, and consequently are controlled only from this page. The plot of the aircraft mission is a two-axis plot that simultaneously displays the altitude profile for the mission as a function of mission time and the Mach number of the aircraft as a function of mission time. This plot allows rapid communication of the flight conditions at each phase of the Rev HK mission. A sample mission is included as Figure 96.

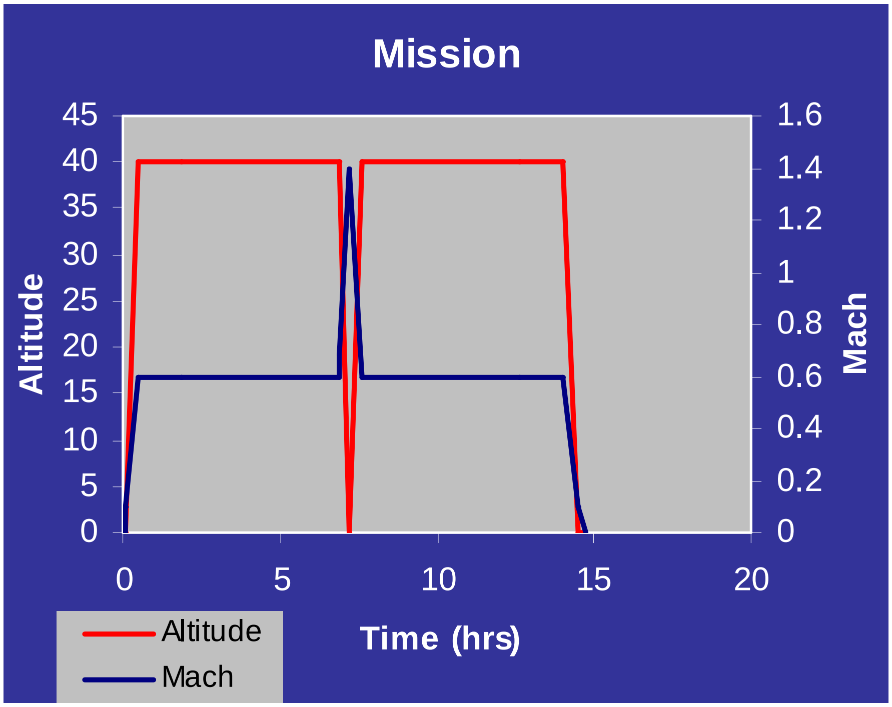


Figure 96: Sample Mission

Environmental parameters allow the selection of the type of atmosphere used for the sizing analysis. Options include Standard, Cold, Hot, and Tropic days, though only the Cold, Hot, and Tropic days are currently available for selection. The selection among these days affects the density conditions calculated for each aircraft mission segment. The selection of atmosphere affects all mission segments; it is not currently possible to assign a different atmosphere to each mission segment.

The aircraft design parameters include the type of engine under consideration and the payload required. The payload value is imported to the sizing module based on the results of the sensor sizing modules and the weapon selection. There are four options for the engine on the Rev HK: reciprocating engine with a propeller, a turboprop, a turbofan, and a turbojet, though data for a reciprocating engine and propeller is not currently incorporated into the sizing tool.

The Rev HK results section is calculated when either the large button above the results section is clicked, or when the "Size Aircraft" button on the main MOA sheet is clicked. The sizing code uses Excel's Solver Function, a numerical optimization function, to converge weights for the aircraft. The Takeoff Gross Weight (TOGW), Fuel Weight, Engine Thrust Required, and Wing Area Required are all calculated and values are returned to the Rev HK results section, and also to the main MOA sheet.

Mission Characteristics Sheet – This sheet is a collection point for values from other places in the code for debug purposes only. The values on this sheet should not be changed.

Segments Sheet – This sheet calculates atmospheric, mission, and other parameters for each segment of the RevHK mission. The segments considered for the mission are warm up, takeoff, climb, cruise, loiter, attack (descend), climb, loiter, cruise, descend, land. Warm up, takeoff, cruise, loiter, descend, and land are calculated as a single mission segment with constant atmospheric, aerodynamic, and engine performance parameters. The two climbing segments and the attack segment are discretized into six sub-segments to account for the variation in atmospheric, aerodynamic, and engine performance parameters associated with the change in altitude. The final descent is considered as a single segment because there is not a speed constraint on the approach to landing, while there is in the attack.

The main areas of calculation for the sheet are Flight Conditions, Air Properties, Installed Thrust Lapse, Drag Polar Components, and Weight Fraction. The values for the aerodynamic K1 and Cdo are currently assumed, but could eventually be linked to a more rigorous aerodynamic module. Weight fractions for the warm up, takeoff, decent, and landing are based on recommendation from Raymer's Aircraft Sizing. [189]

Physics Sheet – This sheet calculates constraints in terms of thrust-to-weight and wing loading. This calculation is based on Equation 16, which is derived from basic force balances of thrust, weight, lift, and drag associated with an aircraft in a steady state. Because of the conceptual nature of this exploration, and the resulting lack of a detailed aircraft geometry for aerodynamic calculations, the entire K2 term for drag in Equation 16 is ignored. Additionally, no drag penalty is considered based on the carriage of stores under the wing or fuselage as opposed to internally.

$$\frac{T_{SL}}{W_{TO}} = \frac{\beta}{\alpha}\left\{\frac{qS}{\beta W_{TO}}\left(K_1\left(\frac{n\beta}{q}\frac{W_{TO}}{S}\right)^2 + K_2\left(\frac{n\beta}{q}\frac{W_{TO}}{S}\right) + C_{D0} + C_{DR}\right) + \frac{P_S}{V}\right\}$$

Equation 16

For each of the mission segments where more than one sub-segment was considered, it was necessary to calculate the constraint line for each of the sub-segments. However, for simplicity on the thrust-to-weight versus wing loading plot, a composite constraint was constructed by using the highest value of thrust-to-weight for each wing loading.

Once all the constraints have been constructed, the "best" aircraft design is that which minimizes the thrust-to-weight ratio for a reasonable wing loading, which was set to have a minimum possible value at 20 lbs/ft^2. In the sizing routine, an optimizer varies the wing loading to obtain a minimum thrust-to-weight ratio, while meeting all of the design constraints. This design point (thrust-to-weight and wing loading) is then used to in conjunction with the weight calculation to determine the required thrust and wing area of the aircraft. A sample constraint analysis is included as Figure 97; in the plot, the abscissa is the wing loading and the ordinate is the thrust-to-weight ratio.

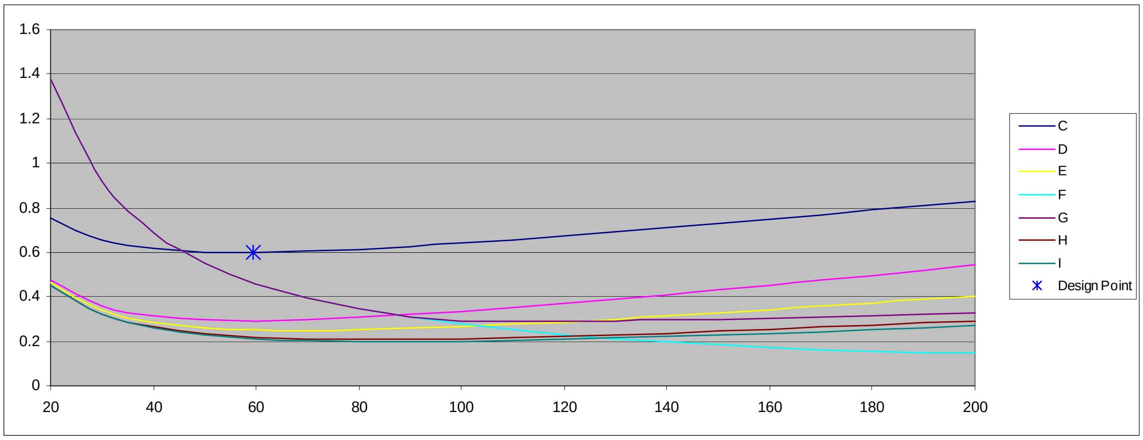


Figure 97: Sample Constraint Plot

Weights Sheet – The weights sheet uses an approach based on Raymer's Aircraft Design to calculate the TOGW of the vehicle (Equation 17) [189]

$$W_{TO} = \alpha W_{empty}^{\beta}$$

Equation 17

The mission fuel fractions from the mission segments sheet are multiplied to obtain the overall mission weight fraction. The values for the mission fuel fraction are calculated based on the Breguet Range and Endurance Equation, or for segments discussed in the mission segments sheet section, using values suggested by Raymer. The mission weight fraction, when multiplied by a TOGW guess and combined with the payload weight can be used to obtain the empty weight of the aircraft. Using the historically based coefficients A and C, a value of TOGW is then calculated. An iterative procedure is then used to reduce the difference between the guessed TOGW and the calculated TOGW. Once the iteration has converged, the TOGW is returned to the main page, along with the fuel weight (based on the mission weight fraction).

Initially, a small database of UAVs was compiled and then regressed to obtain values for the coefficients in Equation 17. However, when attempts were made to re-create a Global Hawk like and Reaper like aircraft, the values of TOGW were much too large. This error was likely because of the small size and fuel capacity of the majority of the UAVs in the database. Because of the errors, the database was replaced with an estimated empty weight fraction based on the Global Hawk and Reaper empty weight fractions. This value is 0.35.

Engine Data Sheet – This sheet calculates the engine thrust lapse (alpha) and TSFC based on Mattingley's historical relationships. The values are representative of engines in each class, but are not tuned for any particular engine. This sheet can be replaced with more accurate engine data when a set of candidate engines are identified. The relationship provided by Mattingley is included as Equation 18. C1 and C2 are empirical coefficients suggested by Mattingley for each type of engine. [153]

$$\alpha_{HBRTF} = \delta_0\left(\frac{p_0}{p}\right)*\left(1-0.49\sqrt{M}\right)$$

$$\alpha_{TJet} = \delta_0\left(\frac{p_0}{p}\right)*\left(1-0.3*\left(\theta_0\left(\frac{\mathrm{T}_0}{\mathrm{T}}\right)-1\right)-0.1*\sqrt{\mathrm{M}}\right)$$

$$\alpha_{\mathrm{Tprop}} = \delta_0\left(\frac{p_0}{p}\right)*\sqrt[4]{(1-0.96*(1-\mathrm{M}))}$$

$$TSFC = (\mathrm{C1}+\mathrm{C2}*\mathrm{M})*\sqrt{\theta\left(\frac{T}{T_{std}}\right)}$$

Equation 18

Air (Table) Sheet – This sheet contains the atmospheric data used in the sizing tool.

5.4.6.2 Economic Analysis Module

The costing calculations for a fleet of UAVs are a difficult task. According to Roskam [194], the total life cycle cost is composed of the planning and conceptual design costs, the preliminary design and systems integration costs, the detail design and development costs, the manufacturing and acquisition costs, the operations and support costs, and, finally, the disposal cost of the aircraft. Roskam breaks the costs down into four areas for estimation: RDT&E costs ($C_{RDT\&E}$), acquisition costs (C_{ACQ}), operating costs (C_{OPS}), and disposal cost (C_{DISP}). Roskam suggests the estimation of these costs largely on a weight-basis, using historical empirical data to establish the relationships. While this approach has worked well for estimating costs of conventional aircraft, for advanced UAVs such as the Revolutionary Hunter-Killer, very few historical data points exist, making actual cost estimation with this approach impossible.

Given the lack of data for costing purposes, the approach for this conceptual study will use cost data for existing UAVs as a baseline, and then modify those baseline costs based on sensor, communication, and weapons characteristics. While this approach will not allow accurate estimates of the hunter-killer costs, applied in a systematic way it will allow the aircraft to be compared. If a higher fidelity study of the aircraft costs is needed, it will be easy to replace the cost estimation module used in this study for one of higher fidelity, and immediately see the impact propagate through to the study results.

5.4.6.3 Sizing and Economic Analysis Results

The system costing environment was constructed in Microsoft Excel ®, and linked to the sizing sheet presented earlier. A baseline mission with 8 hours of loiter time and two 500 nm cruise segments was assumed. For each aircraft concept, the baseline loiter time was adjusted based on the loiter capability of each aircraft in the concept. This meant increasing the loiter time for the Concept 2 killer vehicle by 400 percent, and decreasing the concept 4 loiter times for the hunter and killer aircraft by 20 percent and 30 percent, respectively. The speeds for the baseline mission were not changed among concepts.

Once the mission parameters were adjusted for each concept, the payload weights for the aircraft were specified. The weapon for the sizing was selected based on the AGM-114 HELLFIRE missile [86], with a weight of 100 lbs. The sensor payload was assumed to weigh 20 lbs per unit of coverage. While this assumption is not based on a particular system, it results in sensor weights in the range expected for UAV's performing a search mission. Because sensor technology is difficult to obtain information on and often classified, this assumption removes concern about sensitive data usage. The cost per pound of the aircraft was based on the cost per pound of the Global Hawk. [238] The sizing and cost results are summarized in Figure 98, where each cost is per aircraft. Fleet costs can be obtained by multiplying the average cost by the number of aircraft included in the concept.

	Concept 1	Concept 2	Concept 3	Concept 4	Concept 5
Num HK	1	NA	5	NA	2
Num Hunters	NA	10	NA	15	4
Num Killers	NA	1	NA	8	NA
HK Ws	400	NA	200	NA	200
HK Ww	400	NA	200	NA	400
HK Loiter Factor	1	NA	1	NA	1
Hunter Ws	NA	140	NA	100	200
Hunter Loiter Factor	NA	1	NA	0.8	1
Killer Ws	NA	20	NA	20	NA
Killer Ww	NA	10000	NA	200	NA
Killer Loiter Factor	NA	5	NA	0.7	NA
HK TOGW	18831	NA	16647	NA	17759
HK Fuel Weight	6785	NA	5999	NA	6399
HK Thrust	11342	NA	10027	NA	10697
Hunter TOGW	NA	15124	NA	12053	15484
Hunter Fuel Weight	NA	5450	NA	3917	5580
Hunter Thrust	NA	9109	NA	7260	9326
Killer TOGW	NA	164000	NA	11519	NA
Killer Fuel Weight	NA	C-130	NA	3532	NA
Killer Thrust	NA	C-130	NA	6938	NA
HK Cost ($M)	46.36	NA	42.25	NA	44.36
Hunter Cost ($M)	NA	39.30	NA	33.13	40.01
Killer Cost ($M)	NA	62.44	NA	32.02	NA
Avg Cost ($M)	46.36	41.41	42.25	32.74	41.46

Figure 98: Sizing and Costing Results

5.4.6.4 Terrain and Urban Visibility Modules

5.4.6.4.1 Terrain LOS Calculation

The interference of terrain with sensor line-of-sight is calculated in a MATLAB script. The script currently functions by taking in a 2-dimensional set of terrain values, the altitude of the aircraft, and the width of ½ of the sensor swath. The 2-dimensional terrain values can be easily generated by using the ridgemaker function, or taken as a cross section of the current 3-dimensional terrain over which the aircraft is flying.

The percentage of terrain observable by the aircraft is calculated by discretizing the 2-dimensional ridge into line segments. These line segements are then checked to see if another line segment lies between them and the sensor (masked by a hill), or if they slope away too greatly for the sensor to see their face (a canyon face). An illustration of this is included as Figure 99 where green segments can be seen and red cannot. The area of the total covered area (blue) that can be seen for a given terrain is represented by the green horizontal boxes.

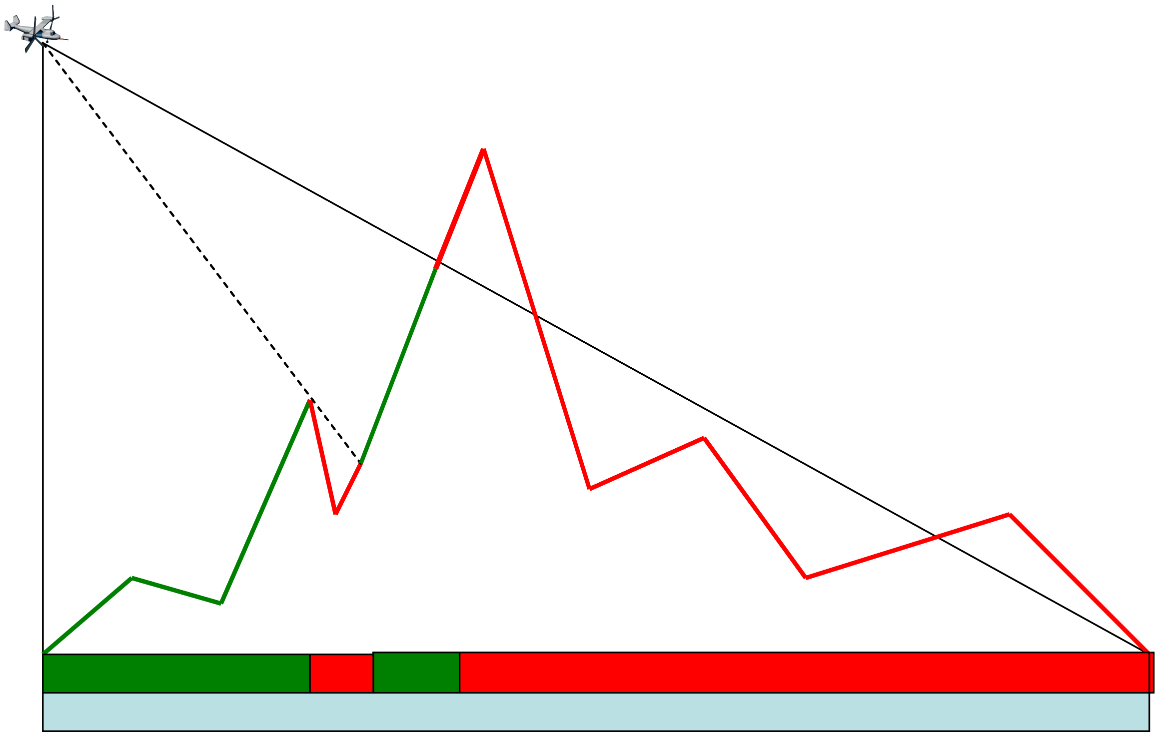


Figure 99: Terrain Masking Geometry

The discretizations are then used to calculate a percentage of terrain visible as a function of the mountainousness, jaggedness, and altitude of the aircraft. The MATLAB script for the ridgemaker and terrain masking calculations are included in Appendix B.

5.4.6.4.2 Urban LOS Calculation

The calculation for the percentage of the streets in an urban area under the aircraft's sensor that can actually be seen by the sensor is calculated as a function of the average block size, the average building height, the average street width, and the altitude of the aircraft. The code assumes that all buildings and blocks are uniform, which is an acceptable assumption since only an average percentage obscured is sought. Additionally, the code assumes that the aircraft is stationed over the center of a building in the center of the block, which is a worst case assumption in terms of visibility.

The urban area is discretized into concentric squares for calculation of the percentage visibility, as shown in Figure 100. The details of the calculation will follow, but as can be seen from the figure, where yellow represents buildings and blue represents streets, the standard grid pattern of a city is not entirely represented by this calculation. The indigo areas of Figure 101 still need to be calculated. As an estimate, these areas are assumed to have the same coverage as the ring before. The aircraft is located at the ‘x’ in Figure 101.

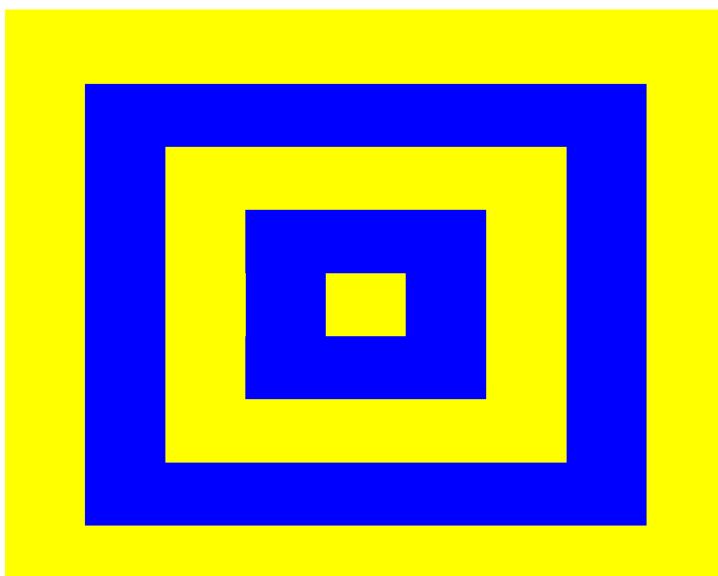


Figure 100: Iteration Scheme for Urban Coverage

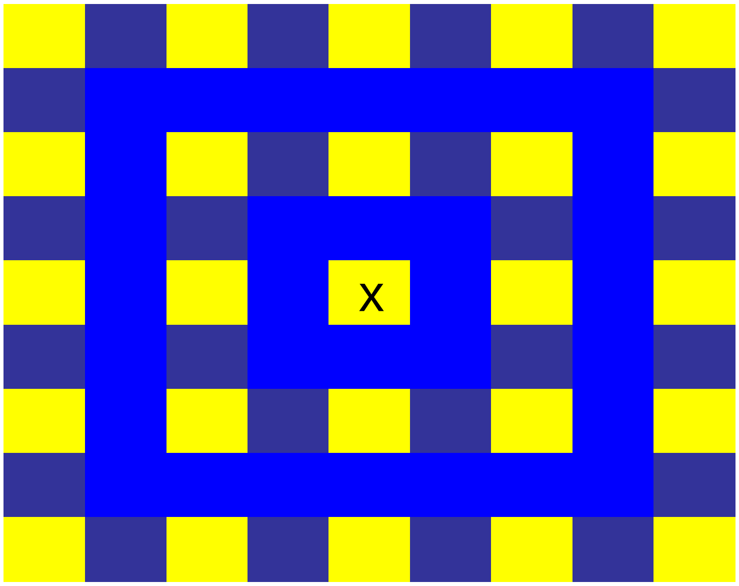


Figure 101: Additional Areas for Estimation in Urban Coverage

Figure 102 shows a 2 dimensional cross section of the sensor coverage, which is assumed to be a square. The aircraft is located over the center of a block, and can "see" the sections of the street that are not blocked by a building. The size of these sections is determined by the height of the buildings, the size of the block, the size of the street, and the altitude of the aircraft. In the figure, the green areas are portions of the street that can be seen, and the red sections are those that are obscured. The size of the red and green sections can be determined from simple trigonometric ratios; right triangles are created by the location of the aircraft, the corners of the buildings and either the top of the building below the aircraft or the ground directly below the aircraft.

The code functions by an iterative procedure from the street nearest the aircraft to the edge of the sensor coverage area (denoted in the figure by the solid cone coming from the aircraft and the blue section at the bottom). In each iteration, the two dimensional areas of the street are then integrated around the blue section in Figure 100. As the integration progresses around the blue section, the distance from the aircraft to the edge of the building grows and shrinks, which is accounted for in the code. Each of the indigo areas in Figure 101 is then estimated based on the percentage available in the closest iteration.

The code then can return several values: the percentage of streets visible, the percentage of the entire swath that is visible street, and the percentage of the swath that is street (both visible and not-visible). The MATLAB script for the urban LOS calculation is included in Appendix B.

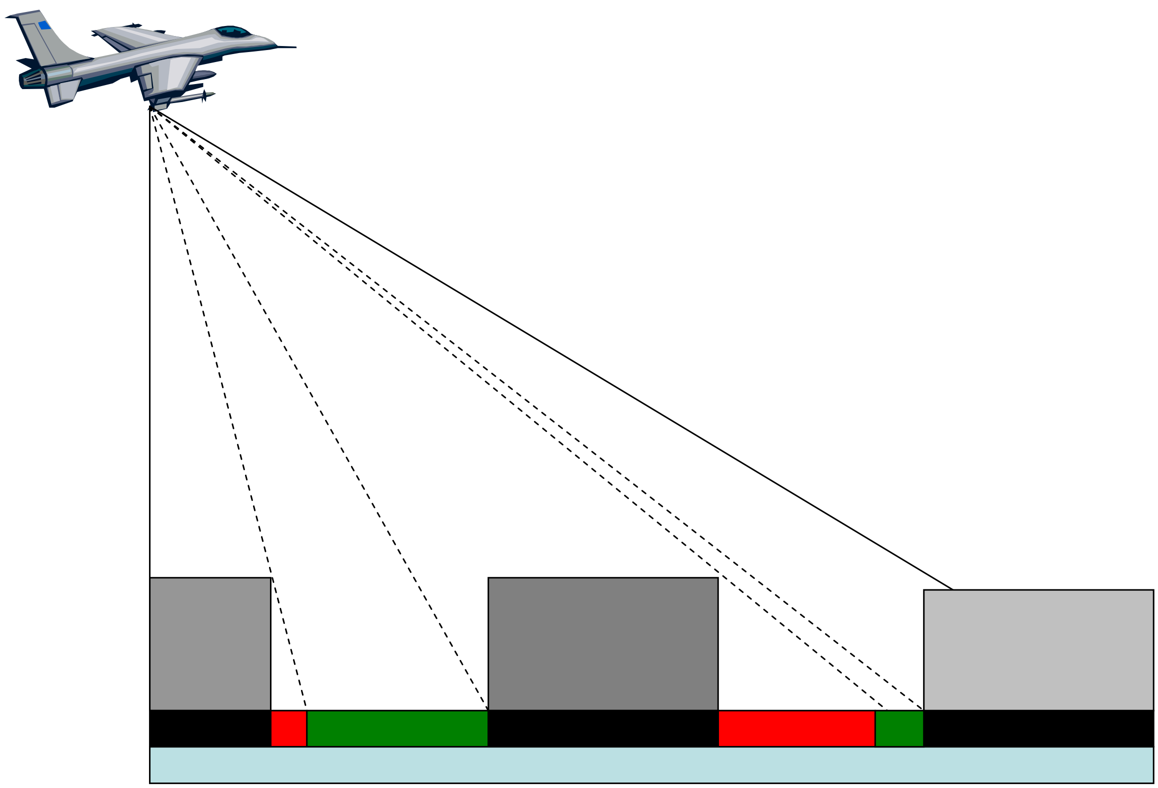


Figure 102: Urban Sensor Coverage Geometry

5.4.6.5 Mission Analysis Module

The mission analysis module operates in a time-stepped fashion, with each agent in the simulation evaluating its location and status at each time step. There are two primary types of agents in the Netlogo environment: patches and turtles. Patches make up the environment in which turtles interact. In general, patches do not move; turtles do. Therefore the environment for the mission analysis module is modeled with patches, while friendly systems, threats, and targets are modeled as turtles. In order to avoid issues with availability of data about hunter-killer concepts, the mission analysis module functions with non-dimensional units that have only properties of length or mass for example, not feet or kilograms. To use the mission analysis module for an actual decision exercise, the appropriate units would be specified and the parametric concepts given appropriate values.

The analysis used a low-intensity search and destroy mission for evaluating hunter-killer concepts. In this mission, time critical targets, fixed facilities, and threats are all present, in a situation analogous to present day operations in Afghanistan. In this situation, long periods are spent searching a relatively large area for targets that do not appear very often. This mission type assumes that little intelligence exists for the direction of the search beyond the general area of interest. The terrain is mixed with small urban pockets that require target masking to be considered. The mission analysis module interface is shown in Figure 103. In the scenario shown in the figure, hunter-killer aircraft are being aided by hunter aircraft, both of which appear in blue. The fixed targets are shown in red, while 1 hiding target (truck) appears in gray. The green and gray patches on the background represent areas of rural or urban terrain, respectively.

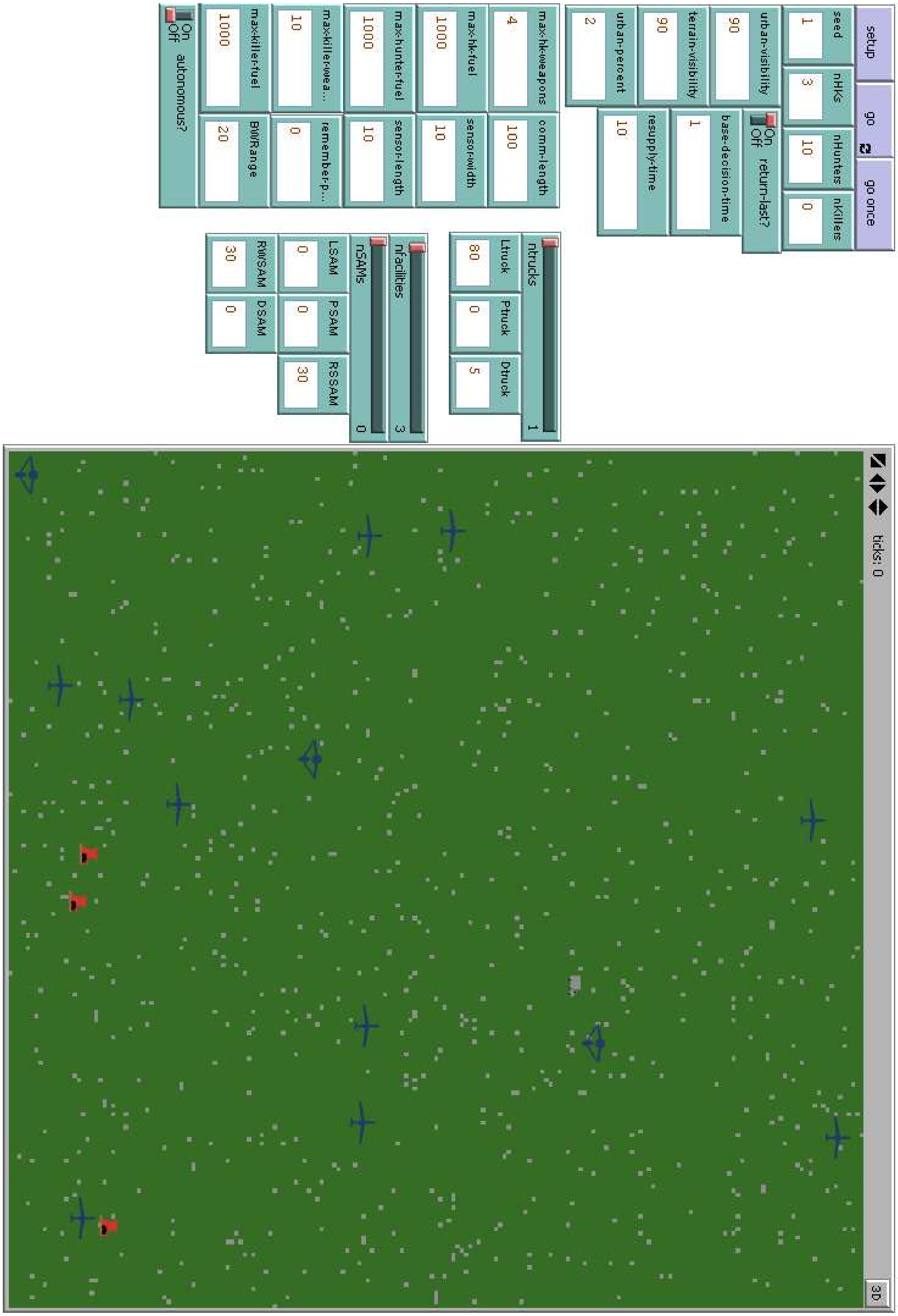


Figure 103: Mission Analysis Module

5.4.6.5.1 Hunter-Killer, Hunter, and Killer Logic

The system-of-systems alternatives were made up of three individual vehicles that were simulated using logic, some of which was shared among the three vehicles and some that was unique to each class. At a high level, the hunter-killer searches for targets in the mission simulation and when it finds one will fire a missile to destroy it. Hunter vehicles search for targets in the simulation, and when they find one they will call for the nearest hunter or hunter-killer to engage the target. Killer vehicles loiter in the area of interest until they are called by a hunter vehicle; when called they fly towards the target and fire a missile at it.

The hunter-killer search pattern is created such that in a cycle through the area of interest, the entire space will be covered by the sensor once. The search path flies the aircraft in the North-South direction as shown in Figure 104. When the aircraft reaches the end of the area of interest, the aircraft moves by the sensor width in the East-West direction, and then flies in the opposite direction of the last sweep. When the aircraft reaches a “corner” of the area of interest, it returns to the start point by the shortest path and starts the search again. The sensor width and length, as well as the aircraft speed, are defined by the user. At each time-tick, the aircraft “fuel” variable is reduced by one to simulate mission fuel burn.

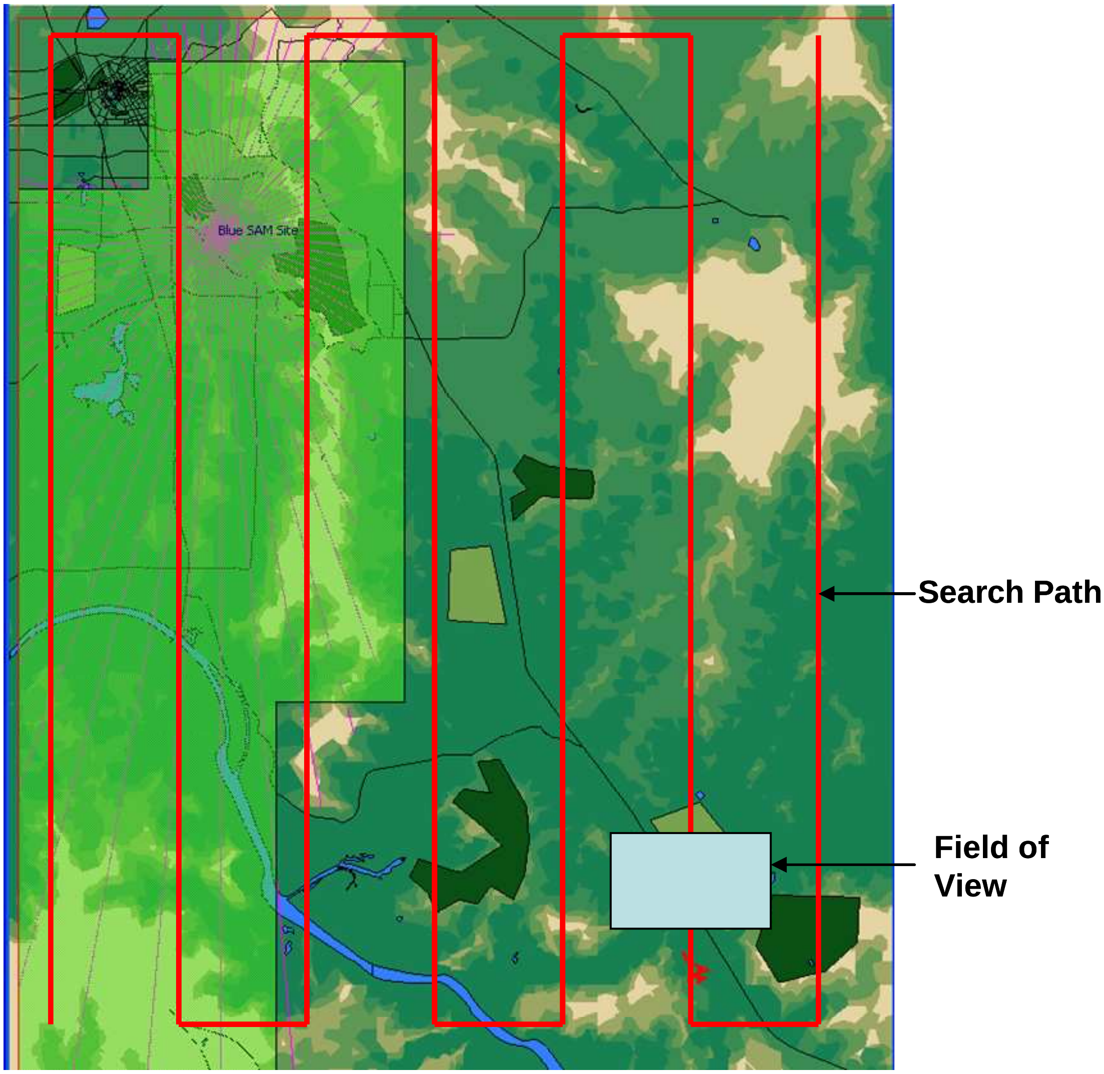


Figure 104: Search Pattern for Hunter-Killer and Hunter Aircraft

At each time step the hunter-killer logic checks the distance to all turtle-targets that are not hiding. If any of the non-hiding targets are within the sensor radius, the aircraft will begin the process of firing on the target. The "decider?" parameter determines whether the hunter-killer is completely autonomous or must ask for permission before firing. If the "decider?" variable is TRUE, the hunter-killer will immediately fire on the target. If the variable is FALSE, the aircraft will wait for the "base decision time," set by the user, plus the discernability factor, which is discussed in the Target Logic section. Once that time has past, the hunter-killer will fire on the target. During the time between the detection of the target and firing, the hunter-killer does not move. This simulates loitering in the area with sensors tracking the target. Once the hunter-killer has fired and successfully destroyed the target, or if the target goes back into hiding before that happens, the hunter-killer will continue with its search pattern.

When the hunter-killer's fuel is reduced to zero (simulating "bingo fuel," when a return to base is required) or has fired all of its weapons, it enters a "resupply" mode. In this mode, the hunter-killer flies directly to the origin of the search pattern and then is "hidden" for a user-specified amount of time. This time simulates either an aerial refueling/rearming, or returning to base for fuel and weapons. Once the resupply time has elapsed, the aircraft flies back to its last search point and resumes searching for targets.

Hunter aircraft follow an identical set of logic to the hunter-killer aircraft, except for when a target is detected. Rather than firing a missile at the target, the hunter calls on a hunter-killer or killer aircraft to destroy the target. The hunter logic compares the distance to the nearest hunter-killer or killer aircraft to a user-specified communication range. If the communication range is greater than the distance, the objective of the hunter-killer or killer aircraft is set to that of the hunter which called them. That aircraft's logic will then direct it to fire on the target, or vector to the target if the distance is too great for an immediate shot.

Killer aircraft follow an identical set of logic to the hunter-killer aircraft, except for their movement. Killer aircraft loiter randomly throughout the area of interest rather than searching for a target. They can only be assigned a target by a hunter aircraft. The random loiter is accomplished by assigning a random heading change between 0 and 15 degrees, and then moving in the forward direction.

5.4.6.5.2 Target Logic

The targets in the mission analysis module are subsets of the turtle class. Three general types of targets are modeled in the NetLogo simulation: trucks, facilities, and threats. The three targets are distinguished by the “variety” parameter of each turtle-target. Each target has the same set of parameters, but the values of those parameters distinguish the target’s parameters.

The simplest form of the turtle-target is the facility variety. Facilities can not hide, have a constant location, and can not defend themselves. The number of facilities is specified in the user interface (or through the batch processing file), and then placed randomly in the environment. The facilities are considered soft, based on the interest in terracotta and dried mud huts for this type of scenario. [254]

The truck variety of turtle-targets operates on a slightly more advanced logic than the facilities. Trucks have four additional parameters that govern their behavior and make it more complex than the facilities. Trucks can hide, and “pop-up” with a frequency specified by a parameter in the user interface, and then will loiter in the “unhidden” mode for a user-specified amount of time. During the time that they are not in hiding, the trucks will move with a user-specified speed in a random fashion. Additionally, trucks have a discernability factor, which is user-specified, and allows a delay to be added before firing to simulate target identification and obtaining clearance to fire in low-intensity conflict.

Threats constitute the final variety of turtle-targets. Threats have “pop-up” and discernability behavior identical to the truck variety of targets, but specified with independent parameters. When not in hiding, threats do not move, and consequently do not have a speed parameter. Threats can, however, fire missiles at hunter-killers, hunters, or killers that come within range when they are not hiding. The detection range and weapon range for the turtle-threats are specified by the user in the interface.

5.4.6.5.3 Missile Logic

Missiles are a turtle-type that is not present when the scenario is first set up. Rather, threats, hunter-killers, and killers “hatch” missiles when they engage another turtle on the adversary’s side. The newly-hatched missiles then become agents within the scenario that follow their own set of logic.

Missiles are assigned the objective of the turtle that fires them. They first check a range to the objective, set by the user, which is the missile’s kill radius. If the distance is less than the kill radius, both the missile and the objective turtle are issued the “die” command. If the distance is more than the kill radius, the missile faces the target and advances based on the missile speed. If the missile “overshoots” the objective in the move, the “die” command is issued to both the objective and the missile as if the missile was within the kill radius.

5.4.6.5.4 Environment Logic

Two types of patches were created for the simulation and are arranged in a grid that is 201 by 201 patches in size. The patch can be defined as either urban or rural. Depending on the type of terrain, a probability is assigned to the sensor detection for the hunter and hunter-killer vehicles. This sensor detection is taken from the urban and terrain LOS calculations. If the patch that the target inhabits is urban, the urban LOS calculation is used, while if the patch is rural, the terrain LOS calculation is used. The ratio of urban to rural terrain is set by a user-defined parameter on the main scenario interface, or in the batch mode input file.

5.4.6.6 Analysis Execution

Before attempting to evaluate the five concepts in the parametric scenario generator, a screening test was conducted to determine the significant factors for the analysis. A Taguchi screening array was used for this test, and 30 cases were executed for each setting to account for the random number generator used in the scenario. The average value of the results was then calculated based on the 30 trials at each setting in the Taguchi array. Based on the results of the screening, eight variables were identified as important for the Global Regret Analysis.

- Number of trucks (time-critical-targets)
- Probability of truck pop-up
- Loiter time of trucks
- Number of facilities (non-time-critical-targets)
- Presence of threats
- Terrain visibility factor
- Urban visibility factor
- Resupply time

For each concept, a DoE was constructed of 12800 cases. The DoE was a combination of a face-centered-central composite design to ensure capture of the corners of the design space and a random sampling of equal size to the central composite design. The two types of the design accounted for 512 cases, which were resampled 25 times each to account for the random number usage in the parametric scenario generator. The ranges for the variables are shown in Table 19.

Table 19: Scenario Generator Variable Ranges

Design Variable	Low Value	High Value
Number of trucks	1	4
Probability of pop up	1	5
Loiter time of trucks	15	150
Number of facilities	0	4
Presence of threats	0	1
Terrain visibility factor	70	100
Urban visibility factor	70	100
Resupply time	10	500

ANN regressions were used to create surrogate models of the data using the same process as described in the previous chapter. All the fits were in the range acceptable for conceptual design; however, the fits for the mission success parameter were typically better than those of the expected friendly attrition rate. The coefficients of determination for each of the ten responses (two responses multiplied by five cases) are included in Table 20.

Table 20: NetLogo ANN Regression Fit Data

Response	Case	R^2	Average R^2 for Response
Mission Success	1	0.9775	0.95616
	2	0.9497	
	3	0.9513	
	4	0.9794	
	5	0.9229	
Aircraft Lost	1	0.8961	0.91278
	2	0.9563	
	3	0.8501	
	4	0.9557	
	5	0.9057	

5.4.7 Compare results

5.4.7.1 Scenario Space Trends

Trends in the scenario space were initially explored using two UTEs, one for each concept's regret and one for each concept's OEC. The two UTEs are shown in Figure 105 and Figure 106, for regret and OEC, respectively. The use of both environments allows the user to explore trends that may be masked by the normalization and comparison that occurs in Global Regret Analysis. However, by using the regret UTE in addition to the OEC UTE, the relative merits of the concepts can be understood in the context of the other systems. The trends shown in the UTEs were created by the surrogate models that were fit to the DoE results.

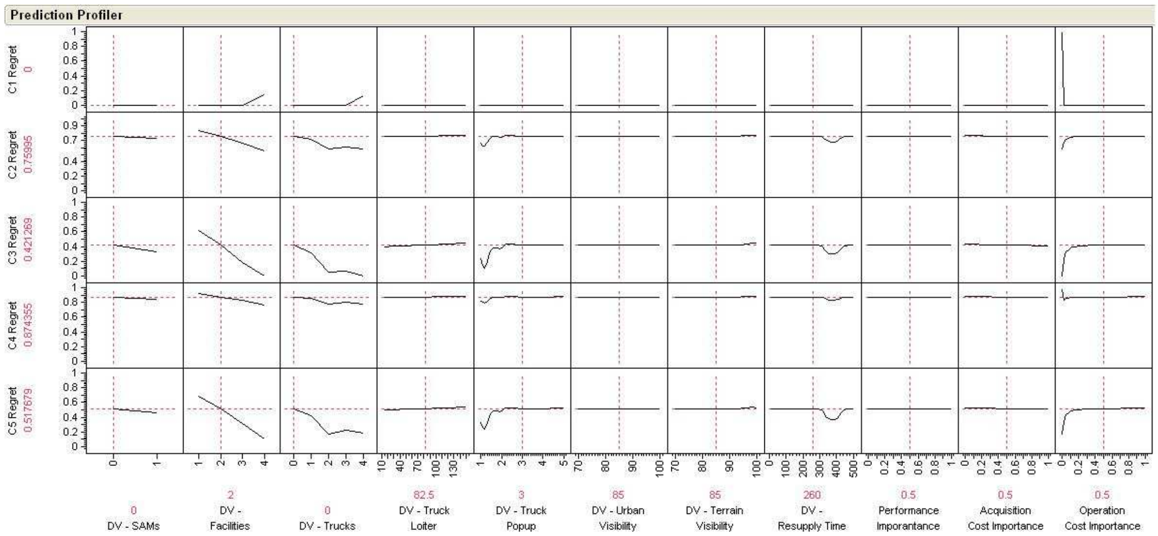


Figure 105: Regret UTE

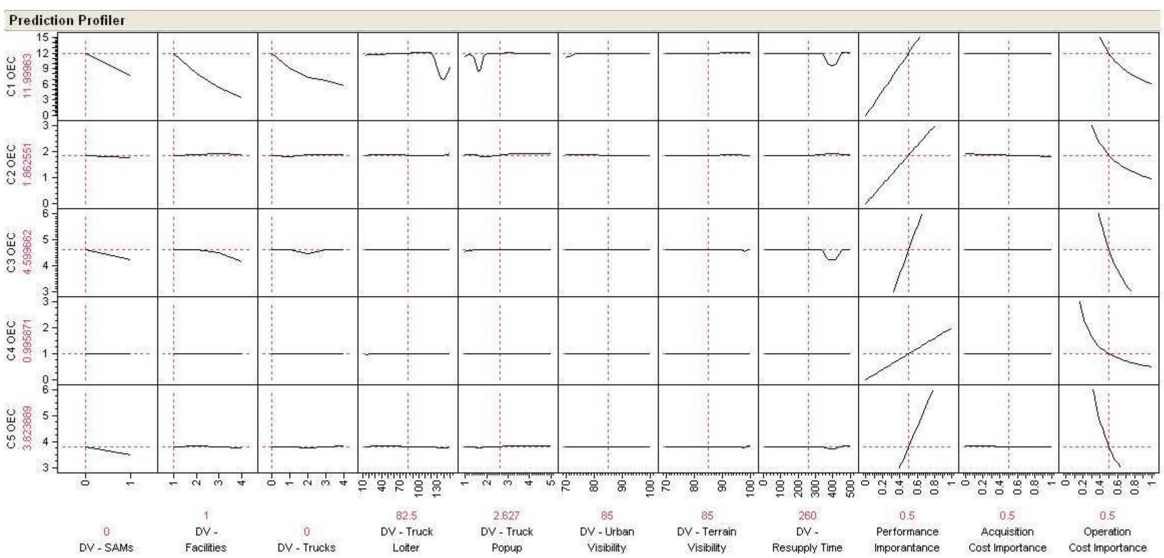


Figure 106: OEC UTE

The following five figures (Figure 107, Figure 108, Figure 109, Figure 110, and Figure 111) show the relative importance and the direction of impact for the eight design variables on the regret associated with each of the five concepts. The most dominant factors change for each of the five candidates. In Figure 107, the first four factors, the number of trucks, SAMs, facilities, and the truck loiter, all result in an increase in regret as the variable increases. This indicates the direction in which the single hunter-killer performance degrades relative to the other concepts. Degradation of performance against an increasing number of targets makes sense for a single aircraft, especially when compared with more dispersed concepts that are able to distribute fires more effectively.

Sorted Parameter Estimates

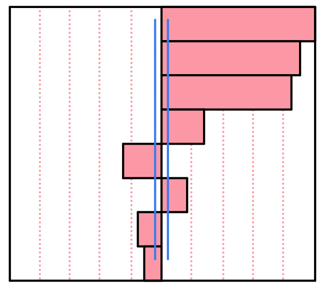
Term	Estimate	Std Error	t Ratio		Prob>|t|
DV - Trucks	0.0418126	0.000825	50.67		0.0000*
DV - SAMs	0.1073234	0.002327	46.12		0.0000*
DV - Facilities	0.0450673	0.001042	43.24		0.0000*
DV - Truck Loiter	0.0004278	2.986e-5	14.33		<.0001*
DV - Truck Popup	-0.012659	0.001008	-12.56		<.0001*
DV - Terrain Visibility	0.0011382	0.000134	8.49		<.0001*
DV - Resupply Time	-0.000064	8.256e-6	-7.76		<.0001*
DV - Urban Visibility	-0.000799	0.000135	-5.93		<.0001*

Figure 107: Concept 1 Regret Factors

The significant factors for the second concept (many hunters and a C-130 type missileer), include the number of facilities and trucks, as well as the truck loiter and popup parameters. The decrease in regret associated with the increase of these factors can be attributable to two trends in the OECs. The "best" concept can be decreasing in fitness or the concept under consideration can be increasing in fitness. In the figure the number of SAMs has almost no effect on the regret of the concept. This is likely because of the distributed nature of the system and indicates resilience to a lost platform.

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
DV - Facilities	-0.045316	0.00039	-116.3	0.0000*
DV - Trucks	-0.019347	0.000309	-62.69	0.0000*
DV - Truck Loiter	-0.000164	1.117e-5	-14.72	<.0001*
DV - Truck Popup	-0.002695	0.000377	-7.15	<.0001*
DV - Resupply Time	-6.316e-6	3.088e-6	-2.05	0.0408*
DV - SAMs	-0.001746	0.00087	-2.01	0.0449*
DV - Urban Visibility	0.0000859	5.035e-5	1.71	0.0880
DV - Terrain Visibility	1.0979e-5	5.015e-5	0.22	0.8267

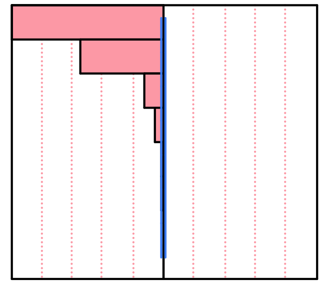


Figure 108: Concept 2 Regret Factors

The number of facilities, trucks, and the presence of threats dominate the variability of the concept with 5 moderately capable hunter-killers similarly to the single hunter-killer. The direction is reversed, however, indicating that an increase in the number of targets reduces the regret for the distributed concept. This is likely a result of the distributed number of vehicles having a combined sensor footprint that is much larger than that of the single hunter-killer. Additionally, the five vehicle concept would not have to rearm as frequently as the single hunter-killer. However, the operating cost for 5 vehicles is significantly more than that of a single vehicle, which is why the regret is decreasing (indicating the dominance of the single vehicle concept). Concepts 3, 4, and 5 all show similar behavior in their dominant factors, supporting the theory that numerous vehicles are more effective as the number of targets and threats increases.

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
DV - Facilities	-0.104033	0.0009	-115.6	0.0000*
DV - Trucks	-0.057851	0.000713	-81.19	0.0000*
DV - SAMs	-0.074901	0.00201	-37.27	<.0001*
DV - Truck Popup	0.0063655	0.00087	7.31	<.0001*
DV - Resupply Time	0.0000195	7.129e-6	2.73	0.0063*
DV - Urban Visibility	0.0002382	0.000116	2.05	0.0405*
DV - Truck Loiter	-4.137e-5	2.579e-5	-1.60	0.1087
DV - Terrain Visibility	-0.000166	0.000116	-1.44	0.1507

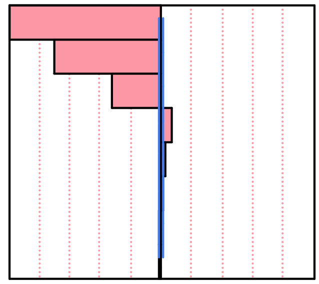


Figure 109: Concept 3 Regret Factors

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
DV - Facilities	-0.024963	0.000198	-126.0	0.0000*
DV - Trucks	-0.013783	0.000157	-87.89	0.0000*
DV - SAMs	-0.027475	0.000442	-62.12	0.0000*
DV - Truck Popup	0.0012682	0.000192	6.62	<.0001*
DV - Urban Visibility	6.4754e-5	2.558e-5	2.53	0.0114*
DV - Truck Loiter	6.8033e-6	5.676e-6	1.20	0.2307
DV - Resupply Time	8.2781e-7	1.569e-6	0.53	0.5978
DV - Terrain Visibility	-1.137e-5	2.548e-5	-0.45	0.6554

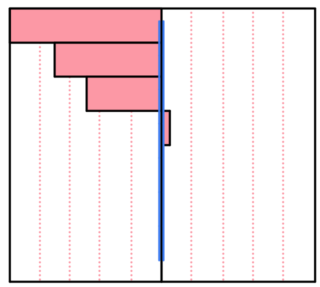


Figure 110: Concept 4 Regret Factors

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
DV - Facilities	-0.091737	0.000744	-123.3	0.0000*
DV - Trucks	-0.044071	0.000589	-74.84	0.0000*
DV - SAMs	-0.030229	0.001661	-18.20	<.0001*
DV - Truck Loiter	-0.0002	2.131e-5	-9.40	<.0001*
DV - Terrain Visibility	-0.000757	9.569e-5	-7.91	<.0001*
DV - Urban Visibility	0.0001593	0.000096	1.66	0.0974
DV - Truck Popup	-0.000762	0.000719	-1.06	0.2892
DV - Resupply Time	1.8532e-6	5.892e-6	0.31	0.7531

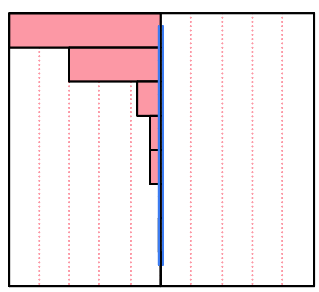


Figure 111: Concept 5 Regret Factors

Figure 112 through Figure 116 give a slightly different perspective on the OEC fore each of the five concepts by including the effect of varying importance of mission success and the two cost factors (coupled with the friendly attrition rate for each concept). In each case, the two most dominant factors are the weight that decision makers give to operation cost and performance. Acquisition cost importance falls into a number of different places in the ranking of factor importance, depending on the concept. Acquisition cost is near the least important for Concept 1(Figure 112) or is the least important for Concept 3 (Figure 114). For the other three concepts it is one of the middle parameters. The importance of the acquisition cost is more likely linked to the attrition rate of aircraft in the various system concepts as opposed to the actual system cost. The system costs were similar across the concepts, but the ones with higher acquisition costs do not necessarily have a higher influence for the importance. Therefore, the most likely candidate is the other multiplier in the OEC that involves the acquisition cost, the attrition rate.

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Operation Cost Importance	-48.60454	5.285928	-9.20	<.0001*
Performance Imporantance	26.423689	5.29403	4.99	<.0001*
DV - Trucks	-3.534441	1.079947	-3.27	0.0011*
DV - Facilities	-2.467868	1.361583	-1.81	0.0700
DV - Truck Popup	1.8789145	1.32103	1.42	0.1550
DV - Terrain Visibility	0.2423507	0.176658	1.37	0.1702
DV - Truck Loiter	-0.024998	0.038924	-0.64	0.5207
DV - Urban Visibility	0.0883372	0.176521	0.50	0.6168
Acquisition Cost Importance	2.6041986	5.246691	0.50	0.6197
DV - SAMs	-1.161429	3.045855	-0.38	0.7030
DV - Resupply Time	0.0010021	0.010814	0.09	0.9262

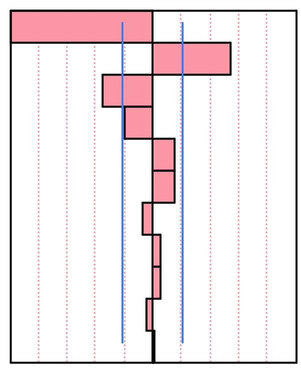


Figure 112: Concept 1 OEC Factors

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Operation Cost Importance	-15.09532	0.694759	-21.73	<.0001*
Performance Imporantance	7.6356207	0.695824	10.97	<.0001*
DV - SAMs	-2.628062	0.400334	-6.56	<.0001*
DV - Resupply Time	0.0042699	0.001421	3.00	0.0027*
Acquisition Cost Importance	-1.453291	0.689602	-2.11	0.0351*
DV - Trucks	-0.257131	0.141944	-1.81	0.0701
DV - Truck Loiter	0.0091813	0.005116	1.79	0.0728
DV - Truck Popup	0.2618925	0.17363	1.51	0.1315
DV - Facilities	0.0819841	0.17896	0.46	0.6469
DV - Terrain Visibility	0.0100805	0.023219	0.43	0.6642
DV - Urban Visibility	-0.007281	0.023201	-0.31	0.7537

Figure 113: Concept 2 OEC Factors

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Operation Cost Importance	-41.83243	2.784811	-15.02	<.0001*
Performance Imporantance	23.614924	2.78908	8.47	<.0001*
DV - SAMs	-6.252874	1.604662	-3.90	<.0001*
DV - Resupply Time	0.0115756	0.005697	2.03	0.0422*
DV - Trucks	-0.807791	0.568954	-1.42	0.1557
DV - Terrain Visibility	-0.054476	0.09307	-0.59	0.5584
DV - Truck Loiter	0.0117825	0.020506	0.57	0.5656
DV - Truck Popup	-0.195272	0.695965	-0.28	0.7790
DV - Facilities	0.1039484	0.717329	0.14	0.8848
DV - Urban Visibility	0.0120875	0.092997	0.13	0.8966
Acquisition Cost Importance	-0.265466	2.76414	-0.10	0.9235

Figure 114: Concept 3 OEC Factors

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Operation Cost Importance	-11.12548	0.595821	-18.67	<.0001*
Performance Imporantance	5.6255878	0.596734	9.43	<.0001*
DV - SAMs	-0.984842	0.343324	-2.87	0.0041*
DV - Resupply Time	0.0033793	0.001219	2.77	0.0056*
DV - Trucks	-0.161055	0.12173	-1.32	0.1859
Acquisition Cost Importance	-0.285619	0.591398	-0.48	0.6291
DV - Facilities	0.0671736	0.153475	0.44	0.6616
DV - Urban Visibility	0.0060895	0.019897	0.31	0.7596
DV - Truck Loiter	0.0011962	0.004387	0.27	0.7851
DV - Terrain Visibility	-0.005154	0.019913	-0.26	0.7958
DV - Truck Popup	-0.001515	0.148904	-0.01	0.9919

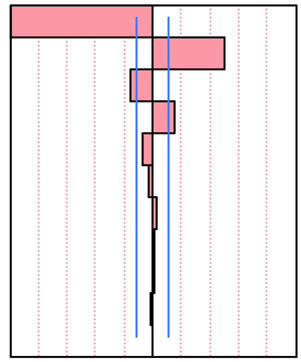


Figure 115: Concept 4 OEC Factors

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Operation Cost Importance	-32.07537	1.64462	-19.50	<.0001*
Performance Imporantance	17.560456	1.647141	10.66	<.0001*
DV - SAMs	-5.467557	0.947662	-5.77	<.0001*
DV - Resupply Time	0.0096299	0.003365	2.86	0.0042*
DV - Trucks	-0.695113	0.336006	-2.07	0.0386*
DV - Truck Loiter	0.0156104	0.01211	1.29	0.1975
Acquisition Cost Importance	-1.479999	1.632412	-0.91	0.3646
DV - Truck Popup	0.2958377	0.411014	0.72	0.4717
DV - Facilities	0.2783912	0.423632	0.66	0.5111
DV - Terrain Visibility	0.014994	0.054964	0.27	0.7850
DV - Urban Visibility	0.0050526	0.054921	0.09	0.9267

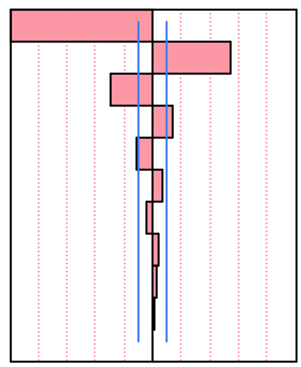


Figure 116: Concept 5 OEC Factors

Figure 117 provides a good example of the visualization of the regret space that is possible with the JMP software. The figure shows the Concept 1's regret as a function of the number of time critical targets in the scenario and the presence of threats in the environment. In places where the surface is flat and equal to zero, the concept is the 'best' choice for the scenario. For the particular settings of the other variables that were used, this area occurs in areas where there are no threats and a relatively low number of targets. However, when more than 3 targets and threats are present, a rapid increase in the regret can be observed.

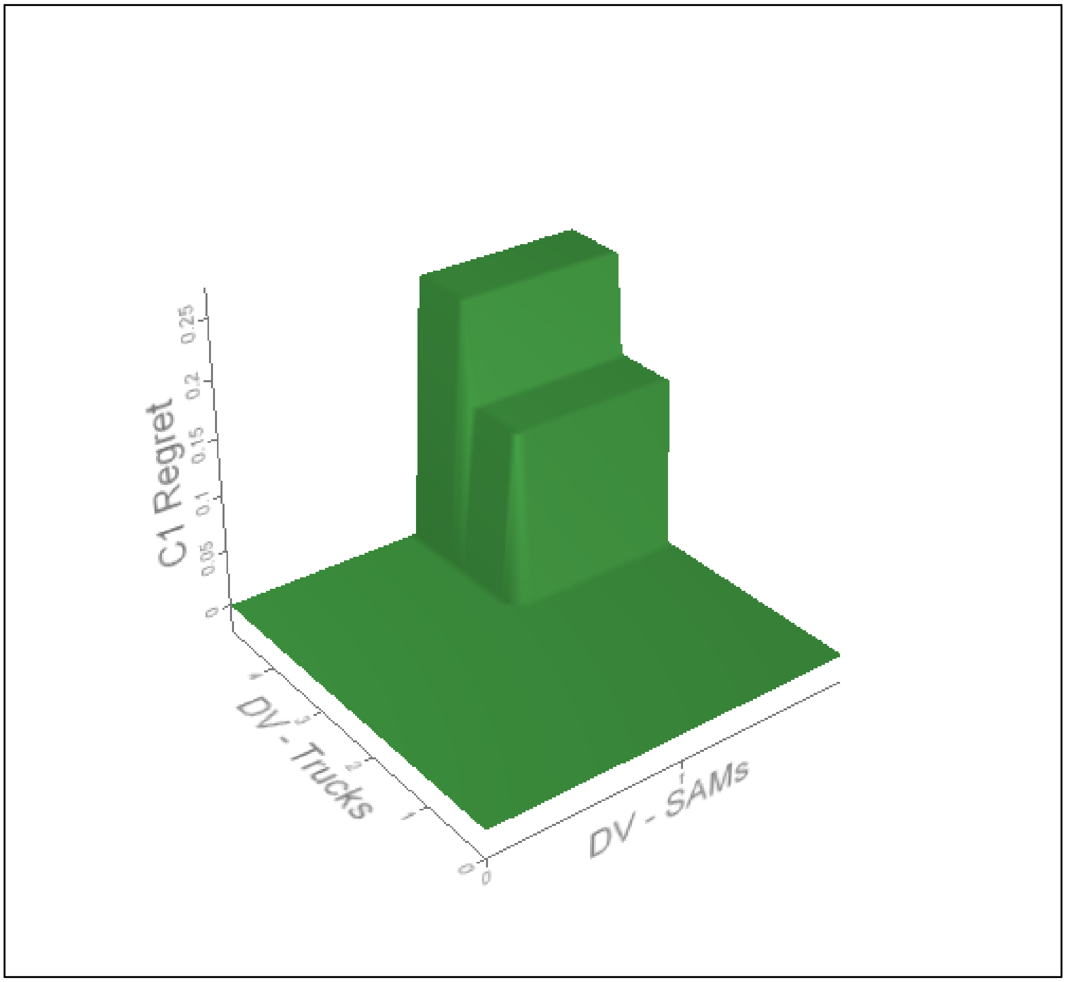


Figure 117: Concept 1 Regret vs. Threats and Trucks

Because there is a rapid increase in regret in one scenario space region of Figure 117, this indicates that a different concept must be the "best" choice for that particular region of the space. Figure 118 shows regret as a function of the same two variables for the other four concepts, but the axes have been flipped so they are unobscured by the surface. As can be seen from the surfaces in the figure, all four concepts display similar responses in regret as a function of the threats and number of time critical targets. However, Concept 3 is the only concept to attain zero regret in the region of high number of time critical targets with threats present. This indicates that Concept 3 is the concept that has overtaken Concept 1 in terms of its OEC. The similarities in the behavior are likely because of the inherent similarities in the concept with respect to using multiple aircraft as opposed to the single aircraft of Concept 1.

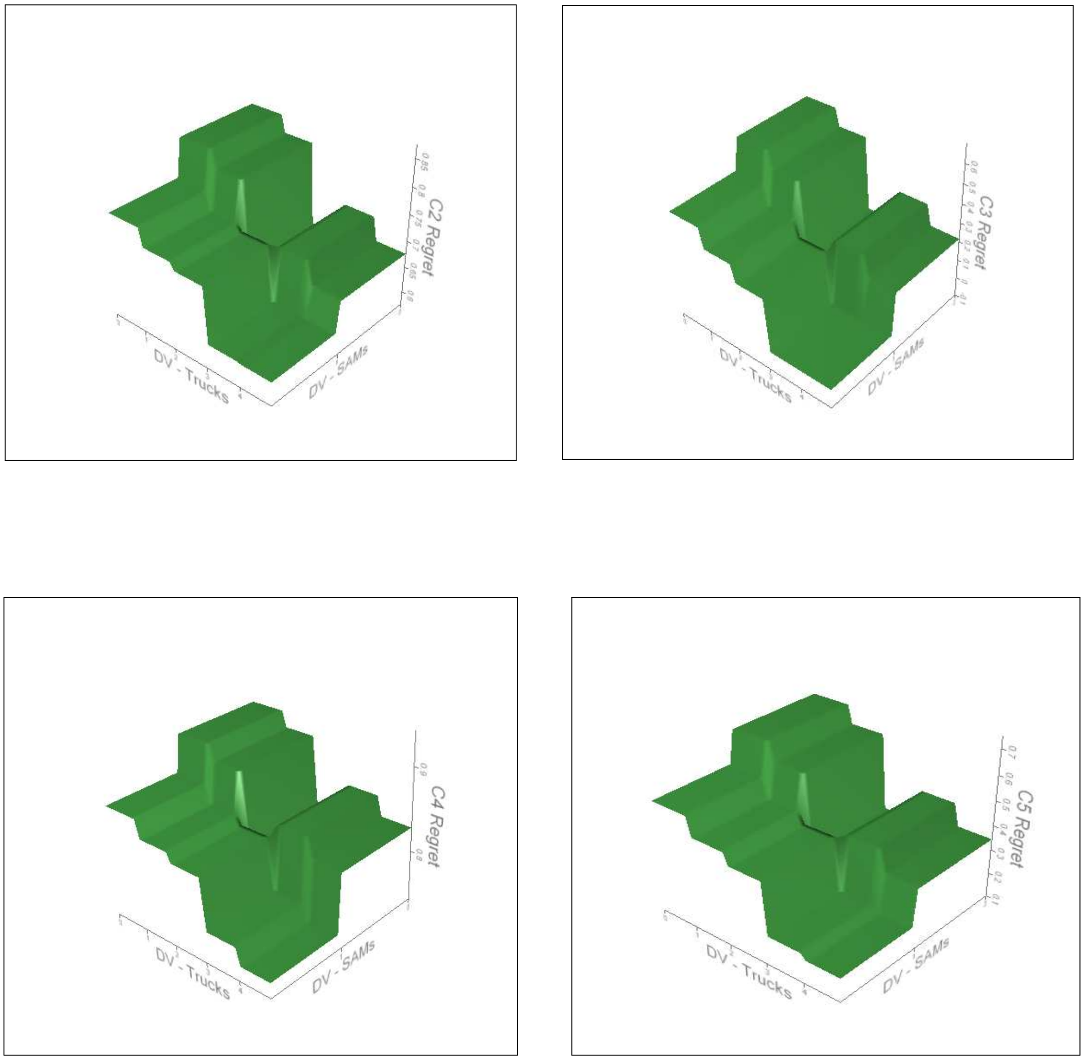


Figure 118: Concepts 2-5 Regret vs. Threats and Trucks

Figure 119 shows a more complex regret response for Concept 1 as a function of the time the time critical targets remain on the field and the time it takes the vehicle to rearm or refuel. The behavior shows the dynamic nature of the scenario space, because while Concept 1 was clearly dominant for the setting of the space shown in Figure 117, there are large regions of Figure 119 where Concept 1 shows regret. There are, however, two regions where Concept 1 is dominant. For long target loiters and long resupply times, and also in one region of resupply time around 200 and loiter around 50. These two areas are mirrored by rises in regret in the other four concepts, shown in Figure 120 and Figure 121.

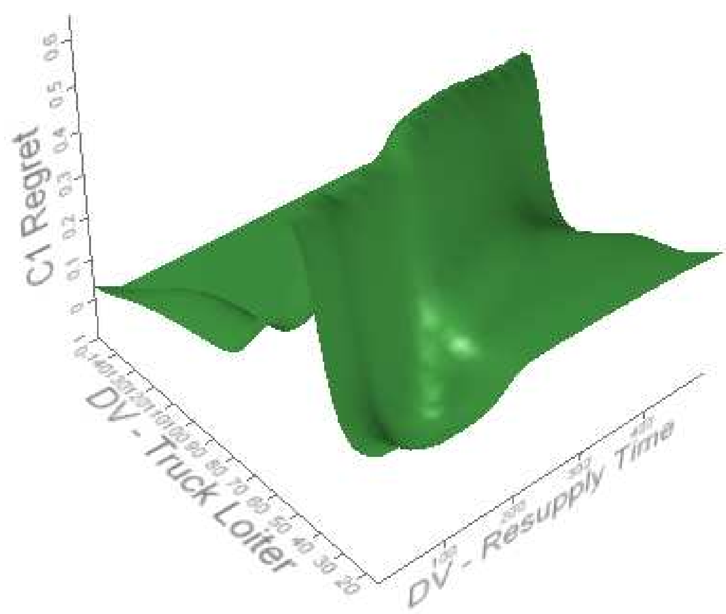


Figure 119: Concept 1 Regret vs. Truck Loiter and Resupply Time

Figure 120 shows the regret for Concept 3 as a function of the time the time critical targets remain on the field and the time it takes the vehicle to rearm or refuel. For the majority of the region of the scenario space shown in the figure, Concept 3 is the dominant solution, showing zero regret. However, for high loiter times, a rise in regret is seen as Concept 1 becomes dominant. The regret for the other three concepts is shown in Figure 121.

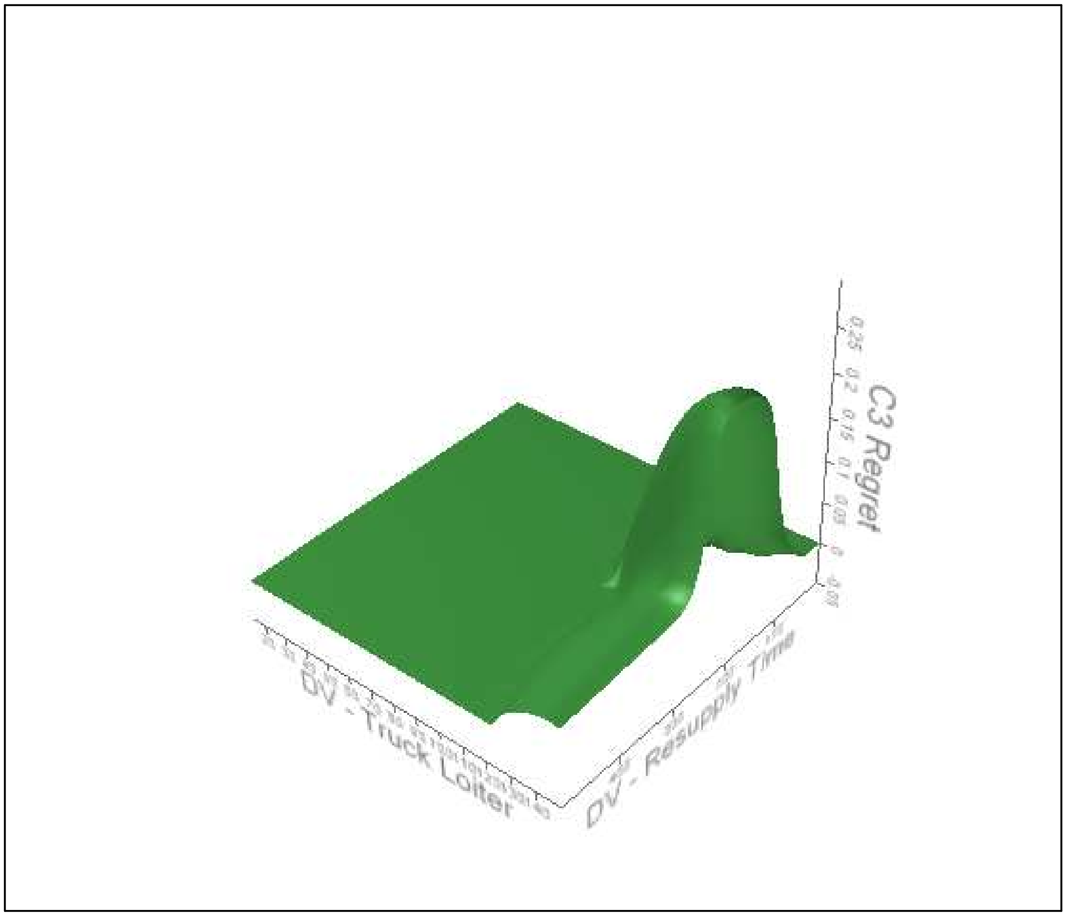


Figure 120: Concept 3 Regret vs. Truck Loiter and Resupply Time

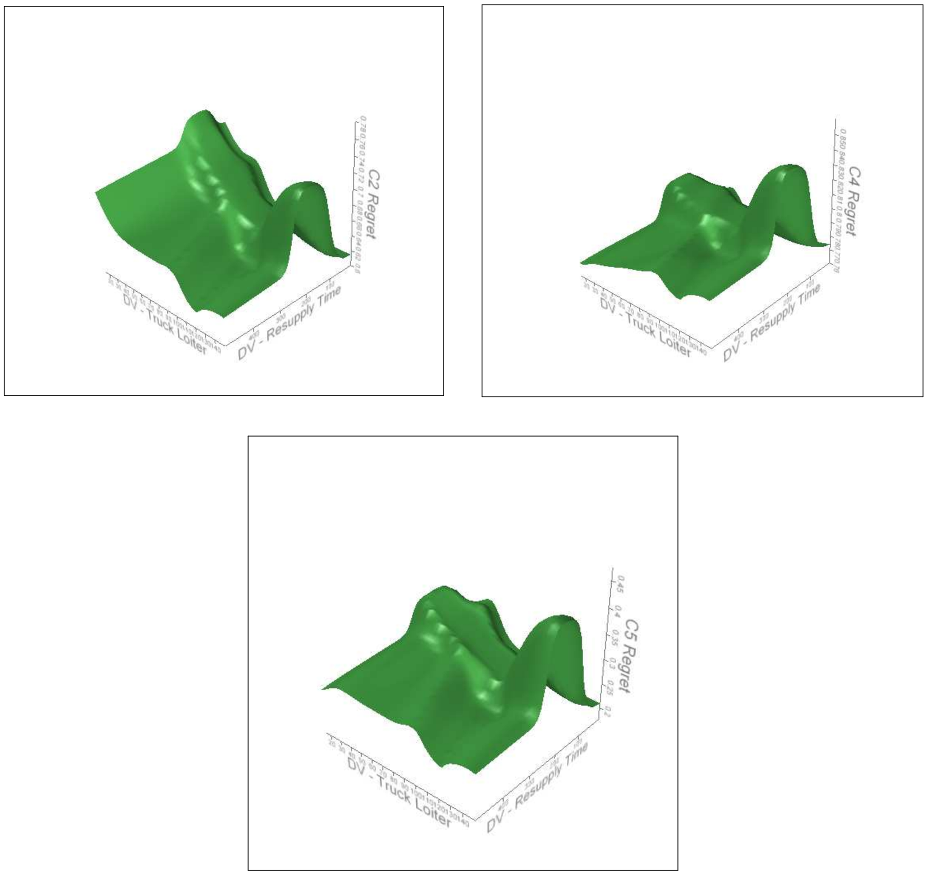


Figure 121: Concepts 2, 4, 5 Regret vs. Truck Loiter and Resupply Time

Another aspect of the regret for Concept 1 is shown in Figure 122. This figure compares regret to the number of fixed targets and the resupply time required. Concept 1 is clearly dominant, except in when there are a large number of facilities and a relatively high resupply time. In this region Concept 3 dominates, as can be seen by the flat region for a large number of facilities in Figure 123. It is interesting to not that there is a region with a large number of facilities where both Concepts 1 and 3 appear to have zero regret. However, close examination of Figure 123 reveals a very slight increase in regret for high numbers of targets but low resupply time. In this region the concepts are very close in terms of their OEC, but Concept 1 has a slight edge. The regret as a function of this aspect of the scenario space for the other three concepts appears in Figure 124.

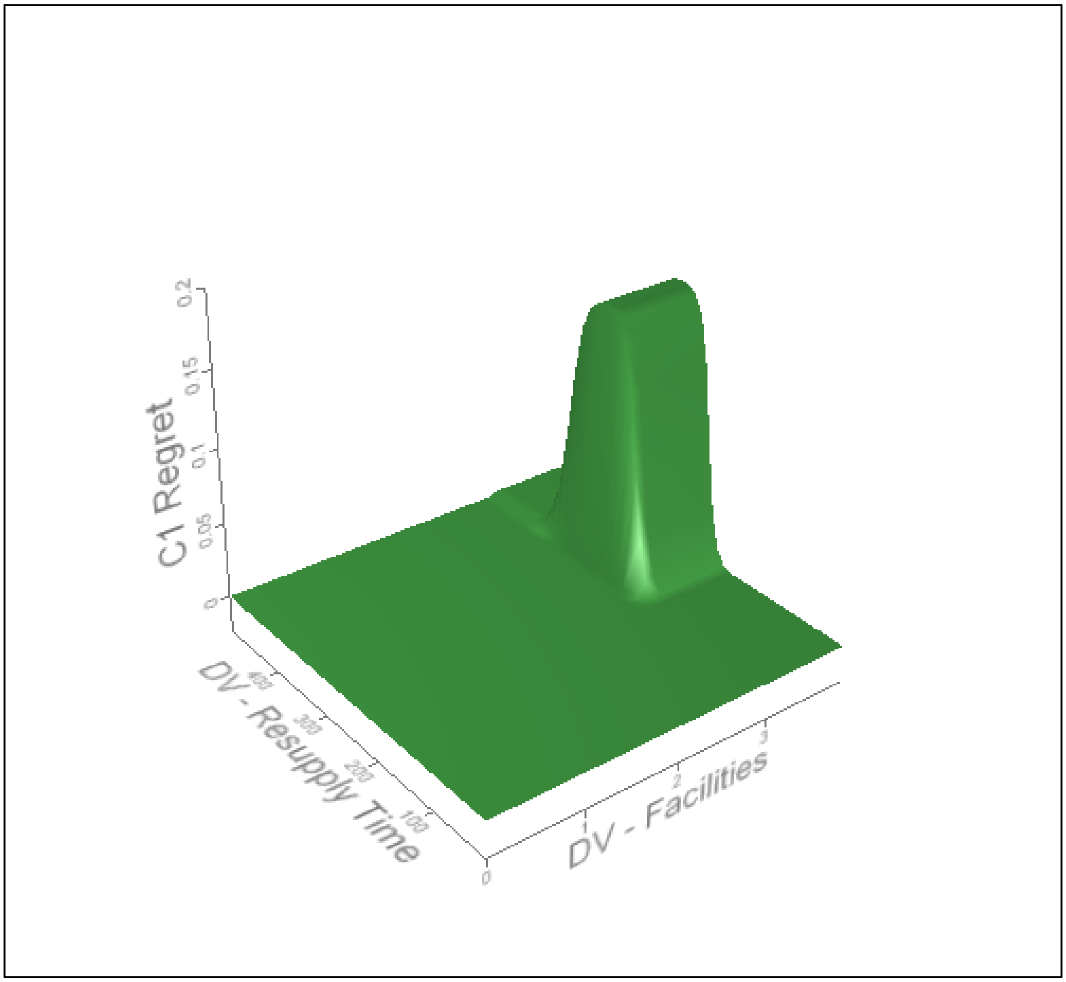


Figure 122: Concept 1 Regret vs. Resupply Time and Facilities

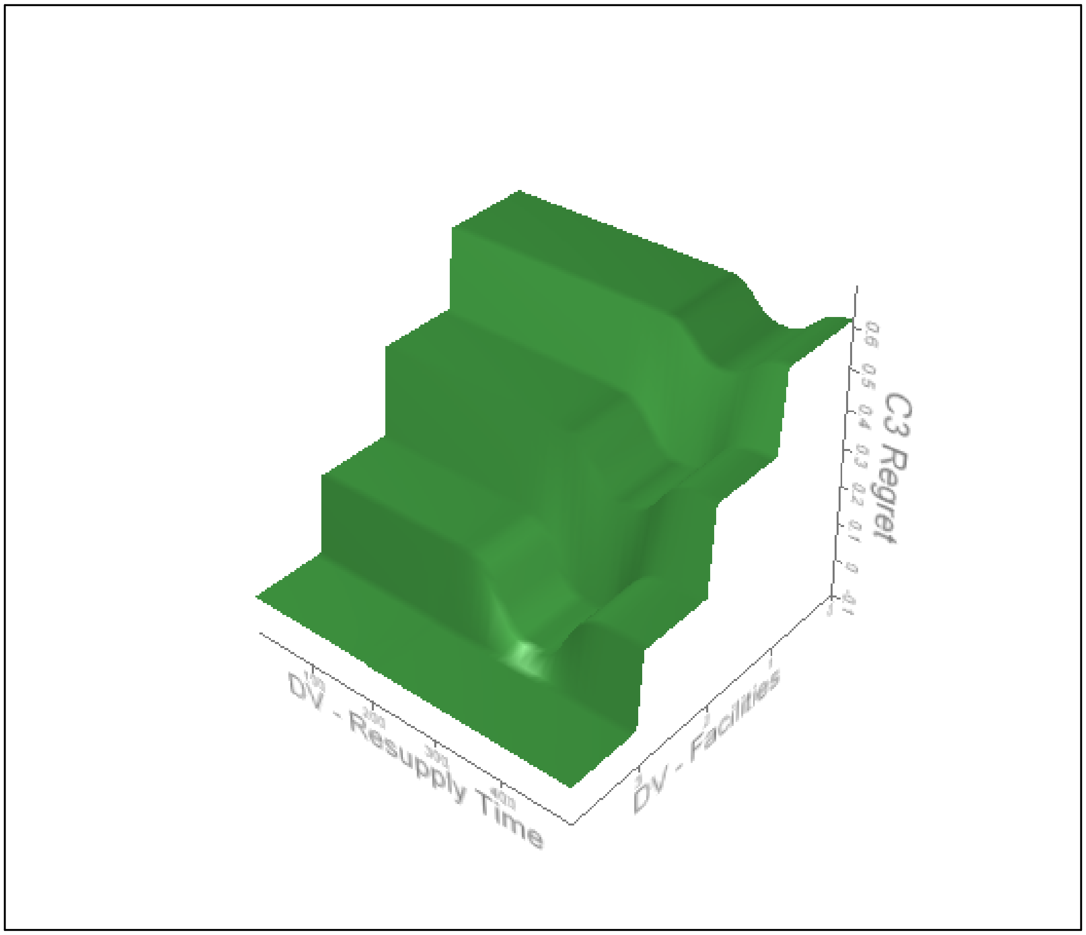


Figure 123: Concept 3 Regret vs. Resupply Time and Facilities

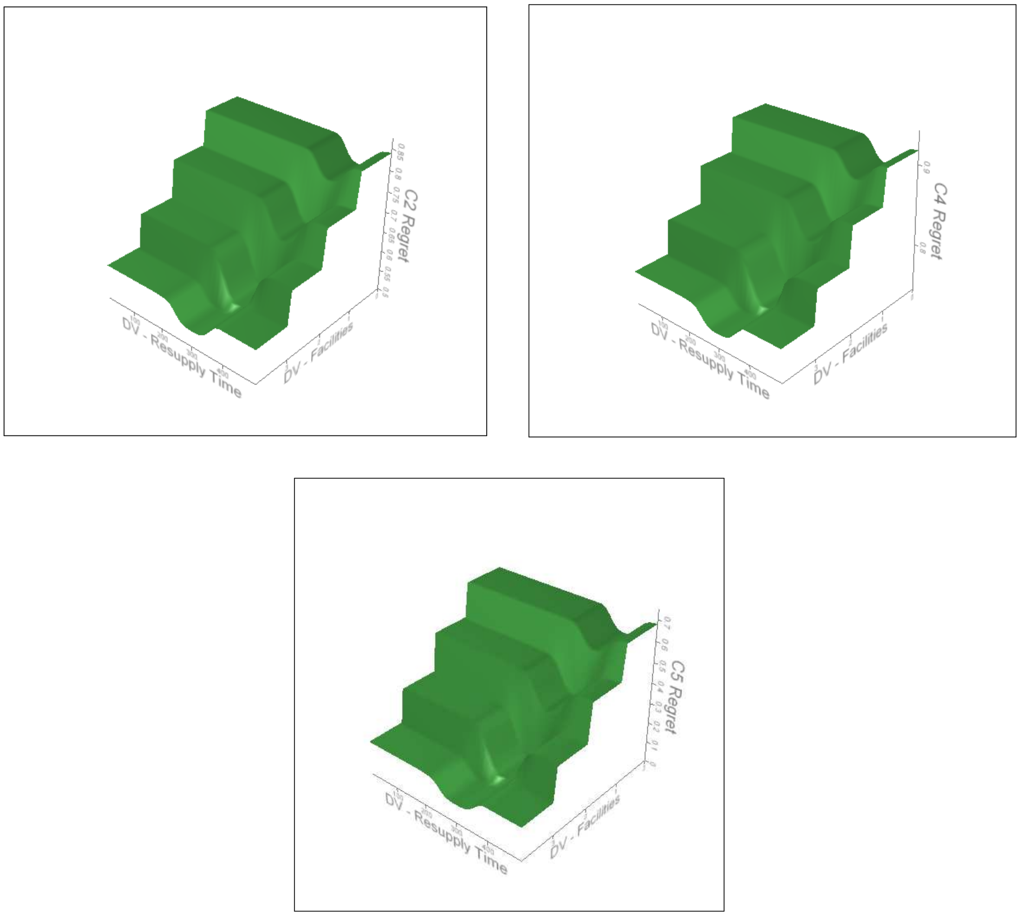


Figure 124: Concepts 2, 4, 5 Regret vs. Resupply Time and Facilities

The ANN regressions used to create the figures shown above were also used to populate a multivariate scatterplot containing the eight design variables, three decision maker factors, and five regret responses. Figure 125 shows the 5000 point multivariate scatterplot. In the scatterplot, the regret responses for the five concepts are shown in the first five rows and columns, the decision maker factors are shown in rows and columns six through eight, and the remaining rows and columns are dedicated to the design variables. Continuous design variables appear as boxes that are "full" of points, while discrete design variables appear with lines of points. There are very few distinguishable trends in the initial population of the design space; however, the addition of color does provide some insight into general trends.

In Figure 125, two regions of the design space have been assigned different colors. Blue points indicate the region of the design space where threats are present and black points indicate non-threatening environments. This color coding allows decision makers to weight the trends observed as a function of two different probabilities of future scenario outcomes.

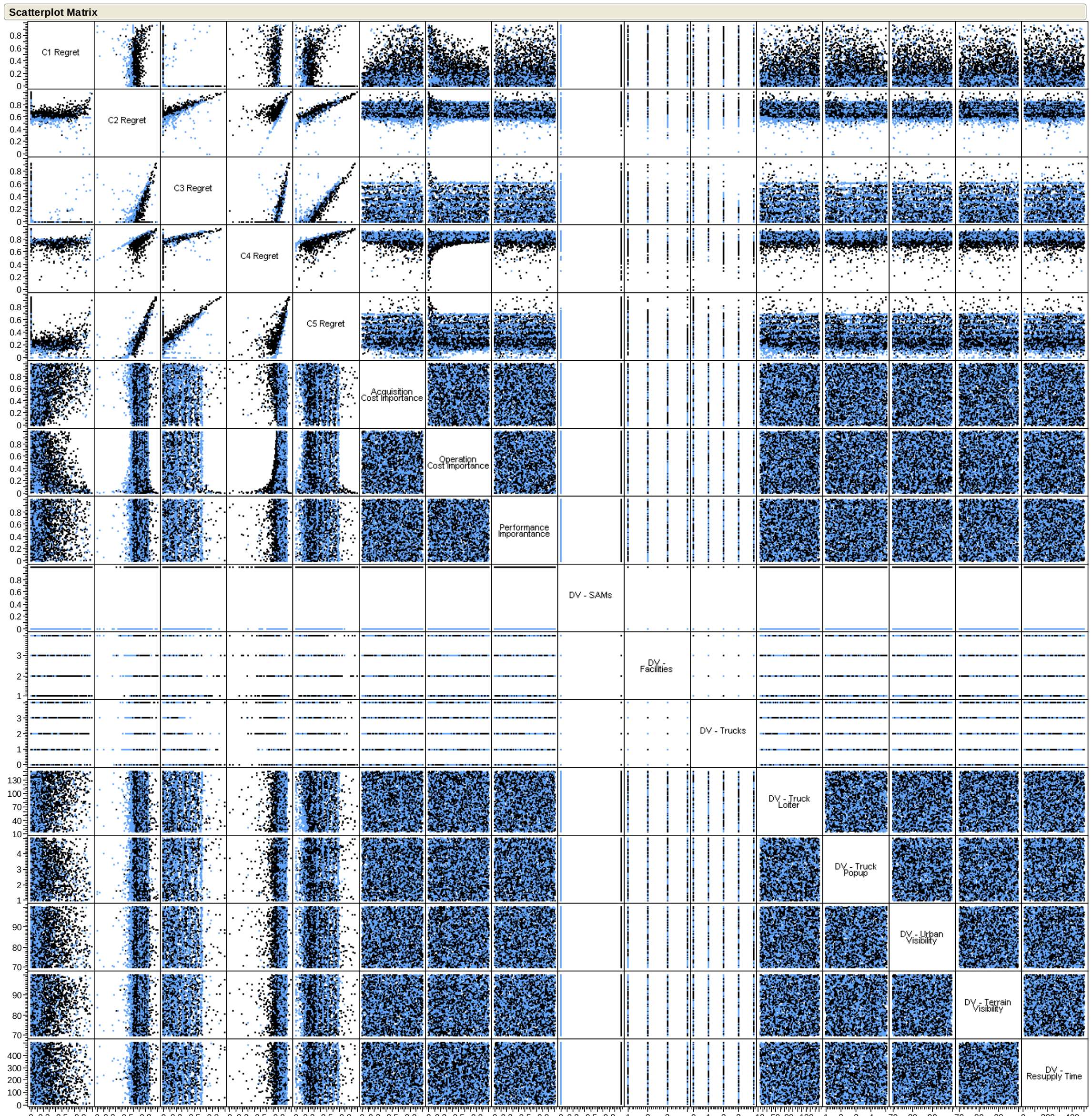


Figure 125: Filtered MCS - Full Scenario Space

Figure 127 shows only the responses and decision maker factors from the full scenario space, using the same color scheme as used above. Some interesting trends emerge in the responses as a function of the decision maker factors. In the first row, sixth and seventh columns, the relation between Concept 1's regret and the acquisition and operation cost importance can be seen. There is a fairly distinct trend where as acquisition cost importance increases, Concept 1's regret increases, and as operation cost importance increases, Concept 1's regret decreases. This result is fairly intuitive as the cost of a single, more capable platform will be higher than a cheaper platform, but the operating cost will be lower than for multiple, cheaper platforms.

Another, very strong trend can be observed for Concept 4's regret with respect to operating cost. As the operating cost importance goes to zero, the regret associated with Concept 4 decreases rapidly, though never quite obtaining a zero regret status. This is because Concept 4 had the highest operating cost of any of the concepts, but included the largest number of vehicles. As the impact of the operating cost decreases, the swarm-effect advantages increase in impact, making the concept more desirable. This trend appears to be more pronounced in the scenarios where threats are present. Because of the distributed nature of Concept 4, threats are less able to destroy the capability of the SoS, increasing its performance in the threatening environment.

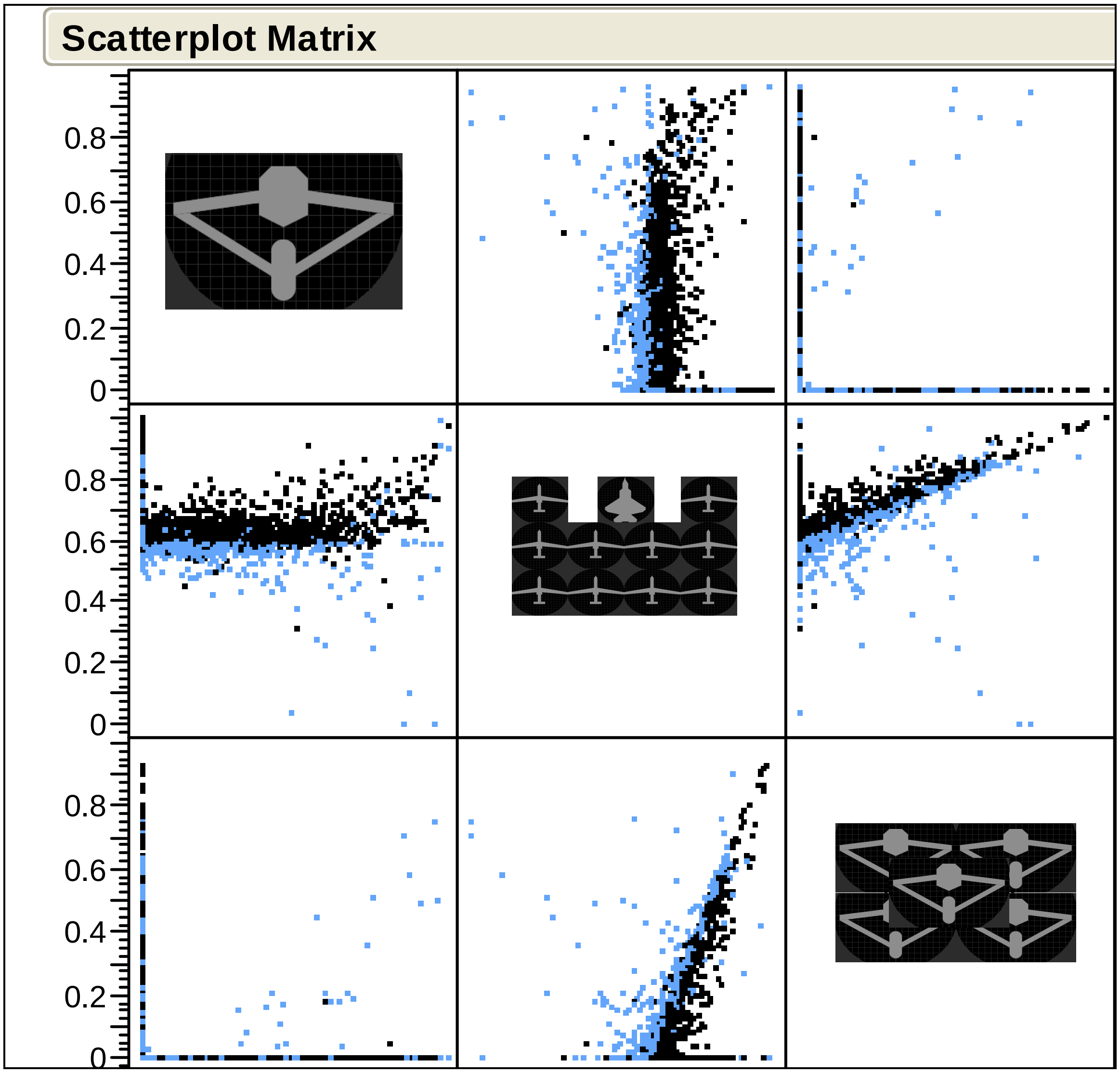


Figure 126: Orthogonality of Local Regret

Figure 126 shows the local regret for three of the concepts considered in the Global Regret Analysis. The top, rightmost and bottom, leftmost plots in the multivariate scatterplot show a relationship among the values of local regret that identifies Concepts 1 and 3 as the dominant solutions for the scenario space. Because the space was populated with a large number of points (5,000), and there are only 26 that are not equal to zero in the horizontal or vertical direction, this means that only 26 points exist where Concepts 1 or 3 are not the minimum-regret solution. The distribution of points in the local regret scatterplot of the two dominant concepts is starkly different than that of the other plots in the multivariate scatterplot. This discovery allows the engineer to quickly identify if a single or pair of concepts dominates the scenario space.

Because so few points lie in the orthogonal set in Figure 126, they can be quickly investigated. The dynamic nature of JMP allows the identification of the scenario variables that led to the other concepts high performance. Once the scenario variables have been identified, the size of the space in which the different concepts dominate can be identified by running a small scenario DoE around the points. The information gained from this type of investigation can be useful for mission planning and understanding the benefits of different approaches in specific regions of the scenario space.

An additional observation may be made about the nature of the points where concepts other than Concept 1 or Concept 3 are "best." In nearly all of these cases, the points are blue, indicating that the scenario environment includes threats. The correlation between other concepts being dominant and the presence of threats in the environment is not surprising, however, because the other concepts have an increased number of aircraft. The increased numbers of aircraft, which are networked, allow the mission to be completed even in the event of several nodes being lost. This is not the case with the single aircraft concept, which, if lost, results in an unsuccessful mission.

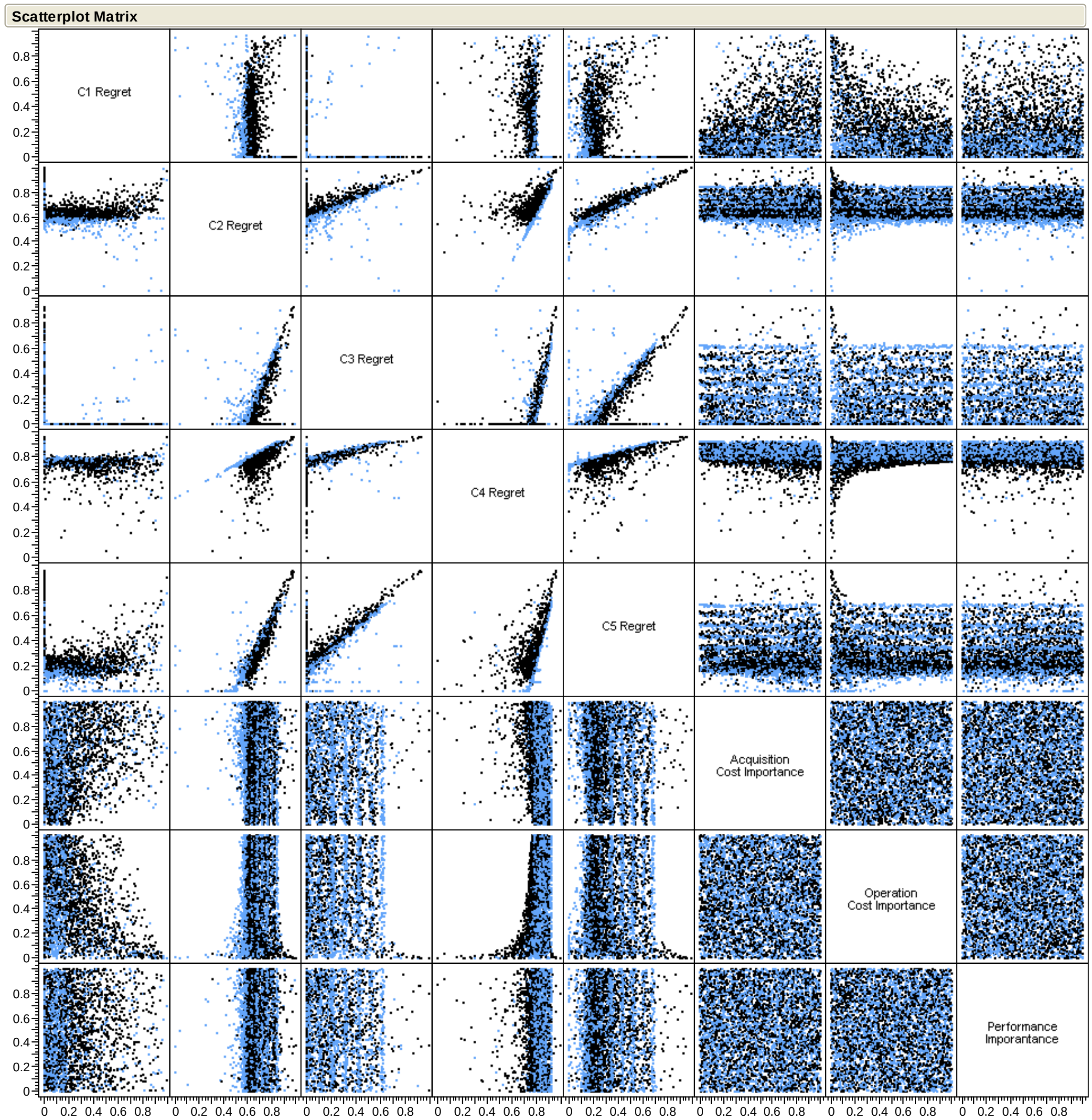


Figure 127: Filtered MCS - Weighting Effects

In Figure 128, an additional filter has been added to the scenario space to simulate the increased time-criticality of the targets. To reduce the space, the bounds of target loiter time were reduced by half (with targets now remaining on the field for half as long) and the likelihood of targets emerging was also reduced. The coloring in the figure indicates the presence of threats in the simulation, with red points indicating threats are present. With this filtering and coloring scheme, Concept 1 clearly has more points with higher regret when threats are present in the environment. This result is likely because of the vast decrease in mission performance that occurs when the only element of the SoS is shot down.

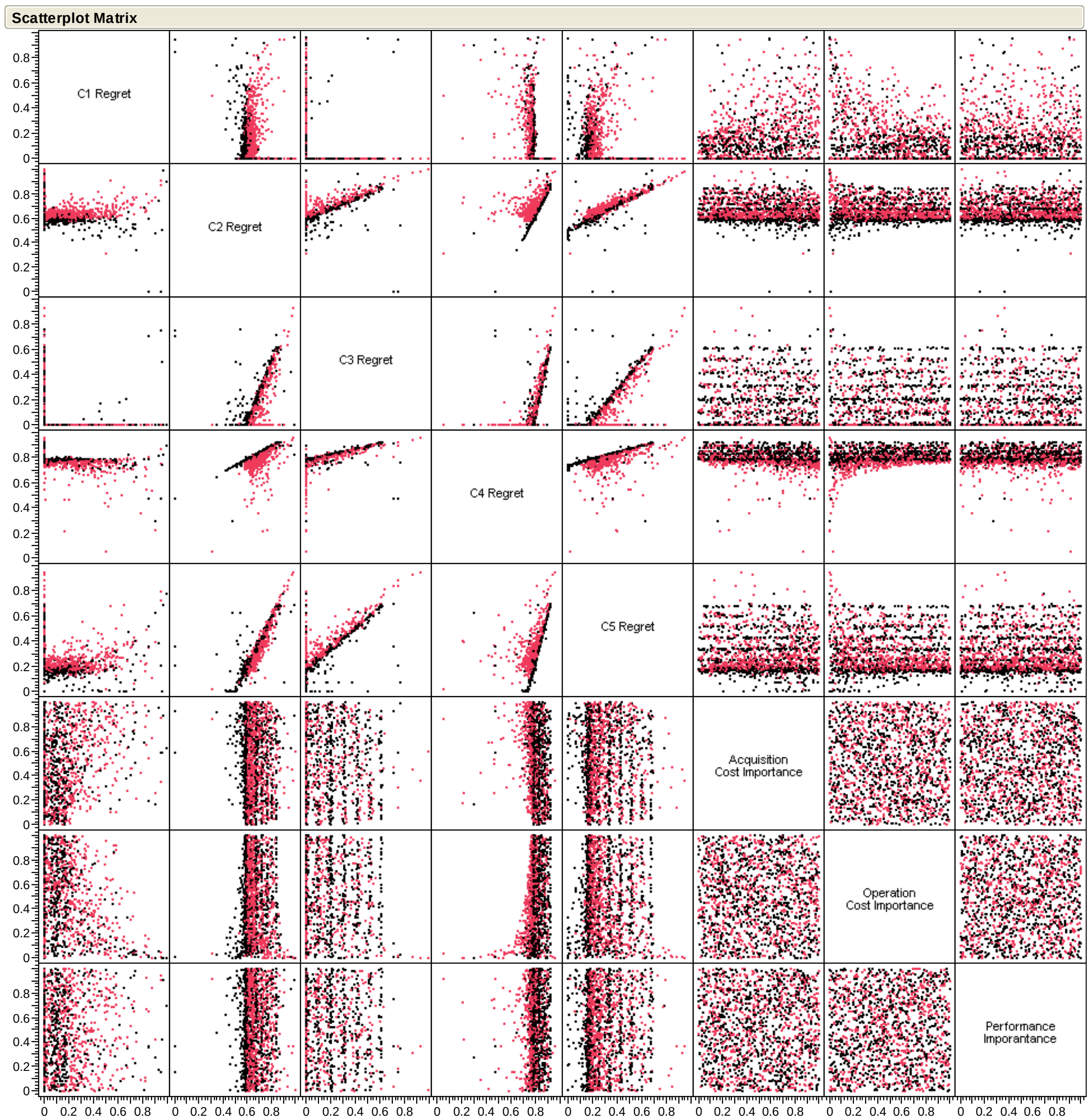
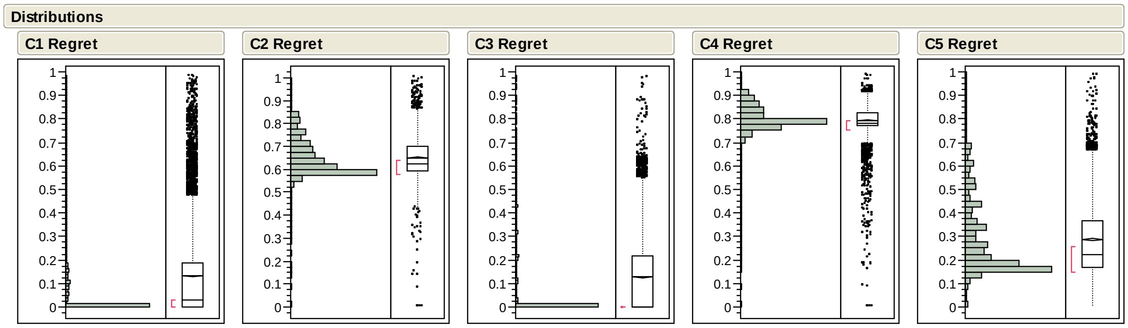


Figure 128: Filtered MCS - Increased Time Criticality

5.4.7.2 Regret Statistics

Ten thousand data points were used to create two sets of statistics for the five concepts. These statistical distributions were created for varying importance of decision maker factors, and constant, equally weighted importance of decision maker factors. Figure 129 shows the distribution data for varying importance factors while Figure 130 contains the results for constant importance factors.

When importance factors are allowed to vary, it is possible for the decision makers to choose a scenario where any concept could be the "best" choice. This can be seen in Figure 129 by the lower bound of each of the five concepts equaling zero. However, Concept 2 and Concept 4 have histograms that indicate the majority of scenarios yield relatively high regrets. This is mirrored in the fact that the 0.5 percent quartile has a positive regret value of greater than 40 percent for both concepts. Concept 5 falls somewhere in between Concepts 2 and 4 and the best two concepts. Its regret histogram shows greatest frequency at a much lower value of regret than Concept 2 and Concept 4, indicating that it is a better choice than those candidates. Concepts 1 and 3 are clearly the best candidates, with their histograms showing the highest frequency with a regret of zero. However, Concept 3 does have a slight advantage over Concept 1, with a mean regret that is 0.004 less than that of Concept 1. This indicates, for the varying factors case, that approximately a half a percent difference between the concepts exists in the OEC on average.



Quantiles

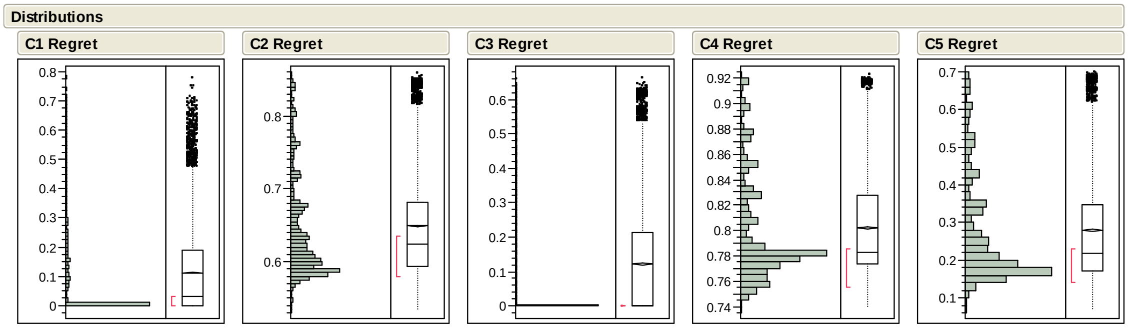
Quantile		C1 Regret	C2 Regret	C3 Regret	C4 Regret	C5 Regret
100.0%	maximum	0.98070	0.99887	0.97594	0.98646	0.98552
99.5%		0.87276	0.89171	0.63436	0.91813	0.77060
97.5%		0.69121	0.83997	0.60135	0.90800	0.67501
90.0%		0.41686	0.77169	0.42994	0.87426	0.53209
75.0%	quartile	0.18838	0.69801	0.21837	0.82831	0.36575
50.0%	median	0.03012	0.62611	0.00000	0.78313	0.22129
25.0%	quartile	0.00000	0.59142	0.00000	0.77307	0.16894
10.0%		0.00000	0.58148	0.00000	0.74758	0.15495
2.5%		0.00000	0.54350	0.00000	0.66666	0.09966
0.5%		0.00000	0.46243	0.00000	0.44933	0.00000
0.0%	minimum	0.00000	0.00000	0.00000	0.00000	0.00000

Moments

	C1 Regret	C2 Regret	C3 Regret	C4 Regret	C5 Regret
Mean	0.1310675	0.6511335	0.1271372	0.7940383	0.2871926
Std Dev	0.1941805	0.0830801	0.1832552	0.066139	0.159117
Std Err Mean	0.0019418	0.0008308	0.0018326	0.0006614	0.0015912
upper 95% Mean	0.1348738	0.6527621	0.1307294	0.7953348	0.2903116
lower 95% Mean	0.1272612	0.649505	0.1235451	0.7927419	0.2840735
N	10000	10000	10000	10000	10000

Figure 129: Regret Statistics - Varying Importance of Factors

Figure 130 presents a slightly different picture than the varying importance factors case. When all of the decision maker factors are given equal weighting, Concepts 2, 4, and 5 retain essentially the same characteristics as in the varying importance factors case. There is one exception to this, however, in that none of these concepts now have a zero regret case. The minimum regrets are 6 percent for Concept 5, 53 percent for Concept 2, and 73 percent for Concept 4. The most interesting observation is that by locking down the decision maker importance factors, Concept 1 shows a lower mean local regret than Concept 3, with a difference of approximately 1 percent.



Quantiles

		C1 Regret	C2 Regret	C3 Regret	C4 Regret	C5 Regret
100.0%	maximum	0.77567	0.85697	0.66007	0.92214	0.69624
99.5%		0.65040	0.84444	0.62075	0.91755	0.68429
97.5%		0.53368	0.82828	0.57186	0.90368	0.65317
90.0%		0.32894	0.76414	0.42320	0.87391	0.52178
75.0%	quartile	0.18923	0.68176	0.21359	0.82796	0.34881
50.0%	median	0.02895	0.62367	0.00000	0.78298	0.21663
25.0%	quartile	0.00000	0.59336	0.00000	0.77312	0.17020
10.0%		0.00000	0.58206	0.00000	0.75950	0.15505
2.5%		0.00000	0.56788	0.00000	0.75130	0.13229
0.5%		0.00000	0.54443	0.00000	0.74729	0.09829
0.0%	minimum	0.00000	0.53279	0.00000	0.73982	0.06628

Moments

	C1 Regret	C2 Regret	C3 Regret	C4 Regret	C5 Regret
Mean	0.1123741	0.6488932	0.1223577	0.8017622	0.2804714
Std Dev	0.151766	0.0721487	0.1771378	0.0425866	0.1461869
Std Err Mean	0.0015177	0.0007215	0.0017714	0.0004259	0.0014619
upper 95% Mean	0.115349	0.6503075	0.12583	0.802597	0.2833369
lower 95% Mean	0.1093992	0.647479	0.1188855	0.8009274	0.2776058
N	10000	10000	10000	10000	10000

Figure 130: Regret Statistics - Constant Importance of Factors

5.4.8 Make a decision

Based on the data presented in the "Compare Results" step of the methodology, the two primary candidates for consideration, based on Global Regret Analysis, are Concept 1 and Concept 3. Regardless of the probabilities associated with time-critical-targets or the threat environment of the scenario, these two concepts were consistently the best alternatives, and the only ones to obtain a local regret score of 0. Based on this analysis, the three concepts that employed dedicated hunter vehicles can be discarded.

Because the Global Regret Analysis changes based on the decision maker weighting factors, it is necessary to carry forward both Concept 1 and Concept 3 to the decision making electronic design review.

5.5 Persistent Strike Evaluation Conclusions

The Global Regret Analysis statistics presented in the previous section provide additional support for the need for exploration of a wide variety of scenarios for comparing alternatives. If any of the 26 cases where Concept 1 or Concept 3 was not the dominant solution was selected as the scenario for comparing alternatives, a choice that is dominated over the vast majority of the scenario space would have been erroneously labeled "best." However, by employing the scenario space exploration using Global Regret Analysis, the dominance of Concept 1 and Concept 3 was clear.

The case explored here also addresses the use of Global Regret Analysis on a true system-of-systems problem. Each of the areas from the Hypothesis Testing section of the dissertation was revisited in the context of this exploration, and no problems arose with the usage of the methodology. The system-of-systems concepts explored in this chapter represented a wide variety of approaches to the problem: single, highly capable vehicles, dispersed roles concepts, swarms of smaller aircraft, and sensor augmentation of aircraft. In the previous pre-conceptual design paradigm, where few scenarios were considered, it was difficult to understand the trades between concepts, especially when concepts performed well under different circumstances. Global Regret Analysis overcomes this challenge by allowing the understanding of the merits and detractors of concepts over the entire scenario space, and then provides a means to weight that scenario space and determine the proper overall judgment.

5.6 Summary

The demonstration of the Global Regret Analysis methodology was conducted using an example problem from the current defense acquisition paradigm. The USAF Persistent, Precision Strike mission provided an excellent opportunity for a relevant, Pre-MS A acquisition program where many SoS alternative exist. All branches of the US military have expressed great interest in UAVs, and the technology provides the possibility to provide significantly increased capability over the current state-of-the-art. The particular program of interest was the Revolutionary Hunter-Killer, and the work for this dissertation was conducted in parallel with efforts at the ASDL at Georgia Tech to support the USAF program.

The general tasks conducted for the evaluation of the alternatives were:

1. Establish the need
2. Define the problem
3. Establish MOPs and MOEs
4. Generate architectures
5. Generate alternatives
6. Analyze alternatives
7. Compare results
8. Make a decision

The majority of the independent work for this dissertation took place surrounding the implementation of the Global Regret Analysis Methodology to the Persistent, Precision Strike mission. Five alternatives were selected from the IRMA for comparison. These concepts represented a wide range of SoS approaches, including single-vehicle approaches, teams of similar vehicles, teams of different vehicles, and swarm concepts.

The Parametric Scenario Generation M&S environment was constructed in MATLAB and NETLOGO, and linked using ModelCenter. MATLAB was used to construct terrain generation, urban visibility, and terrain visibility models. NETLOGO was used to evaluate the effectiveness of the different concepts in a time-critical target prosecution mission. The simulation included the ability to manipulate target characteristics, threat characteristics, and terrain. All of the simulation software was written by the author of this dissertation.

The effectiveness results of the Parametric Scenario Generation M&S were evaluated in the JMP statistical discovery environment. Regret analysis showed interesting trends in the scenario space and identified two concepts that dominated the vast majority of the scenario space. However, cases could be found where each of the concepts would be dominant, underscoring the need for robustness evaluation when considering alternatives in the defense acquisition process.

CHAPTER 6

CONCLUSIONS

6.1 Final Experiment

This dissertation was undertaken with the hope of improving the ability of the early defense acquisition process to understand and account for robustness in the design of military systems and systems-of-systems. The question then becomes, because defense systems can only truly be evaluated after their service life has ended [183], and because defense systems are too expensive to provide control and experimental alternatives, how does one determine if the new way is better than the old.

While modeling and simulation approaches have shown that using Global Regret Analysis can identify systems that perform “better” across many possible manifestations of friendly tactics, environmental conditions, and enemy tactics and technologies, the performance of the system in the “real world” is much more complex. In all likelihood, the system under consideration will only have to perform in a few conflicts; the billions of possible futures will only manifest to a handful. In reality, we can never know if the current predictions of what the face of war will look like in the mid-to-long-term will come true, or if a completely different and unexpected paradigm of warfare will emerge. Therefore, the only way to understand the impact of this dissertation is to document the decisions that “would have been” during the electronic design reviews and, many years in the future, assess how the decisions that were made compare.

Because of this inability to validate through experimentation the strength of the methodology presented in this dissertation, a more qualitative approach must be taken if any near-term understanding is to be obtained. In the earliest phases of the literature search, a qualitative comparison was made among the current methods for assessing robustness in design. This search was based on the results of a thought-experiment, an exercise in logic, which identified an improvement in robustness as desirable for military systems. Because we strive to understand how the current paradigm compares to the work of this dissertation, the same qualitative comparison exercise will be used to try and understand how the new method compares with the state-of-the-art.

Figure 131 shows the initial assessment that was used to base-line the existing robustness evaluation techniques (repeated from Chapter 1). The criteria in the left-most column of the figure will now be used to provide an assessment of the new method. The importance of the various criteria will then be manipulated, and using the Technique for Ordered Preference by Similarity to an Ideal Solution (See Appendix A), the robustness of the new method to various weighting schemes will be explored.

Poor - ●
Moderate - ⊙
Good - ○

	Not Optimizer Based			Optimization	
	Kazmer and Roser (1999)	Ford and Barkan's Robust Concept Design (1995)	Taguchi's Parameter Design	Lewis and Parkinson (1994)	Wilde (1992)
Mathematical Definition of Robustness	○	○	⊙	⊙	○
Applicability in Conceptual Design	⊙	○	⊙	●	⊙
Applicability in Pre-Conceptual Design	●	⊙	●	●	●
Robustness Evaluation at Capability Level	●	●	●	●	●
Applicable to Systems-of-Systems	⊙	⊙	⊙	⊙	●
Applicable to Multi-Objective Problems	○	⊙	⊙	⊙	●
Optimizable	○	⊙	⊙	○	○
Automated	●	●	●	○	○
Applicability to Revolutionary Concepts	●	⊙	●	●	●
Robustness Evaluation Based on Full Life Cycle	⊙	○	●	⊙	●

Figure 131: State-of-the-art Robustness Evaluation Techniques

6.1.1 Mathematical Definition of Robustness

The Global Regret Analysis approach does provide a concise, mathematical definition of robustness. This mathematical definition is the integral, over the entire possible futures space, of the probability weighted local regret values.

This definition, however, does not come without challenges. In order to successfully evaluate the Global Regret, the designer must be able to integrate the probability weighted local regret function over all the dimensions of the possible future space. This integration is not a trivial task, and can require significant computational power. The Global Regret can be approximated, however, by considering the mean local regret of a large sampling of the possible futures space, and the distribution of the local regret in the possible futures space.

An additional challenge associated with the mathematical definition of robustness as Global Regret arises from the discrete nature of traditional regret analysis. To evaluate the Global Regret, a continuous function must be created for the response data. However, the use of highly accurate surrogate models has been shown to overcome this challenge.

6.1.2 Applicability in Conceptual Design

The Global Regret Analysis approach can be used in conceptual design, assuming quantifiable data on the measures of merit of interest can be created. Additionally, the designer must be able to identify the scenarios over which the system or SoS will be expected to be used. An additional, useful piece of information would be an estimation of the likelihood of the different scenarios.

In a modern systems and SoS design environment, the conceptual design phase incorporates physics-based modeling, historical and empirical relationships, and capability analysis. These models are created and integrated in such a fashion as to return quantification of the different concepts with respect to the measures-of-merit for the program. However, these models may not be valid over the entire scenario space. As a result, if the designer wishes to gain understanding of the robustness of particular concepts relative to the other design alternatives by using Global Regret Analysis, particular attention must be paid to the suitability of models over the entire scenario space.

6.1.3 Applicability in Pre-Conceptual Design

Applicability of Global Regret Analysis in the pre-conceptual design phase is dependent on the availability of data quantifying the measures of merit for different approach alternatives. In the pre-conceptual design phase, decisions are often made on the basis of heuristics, expert surveys, and back-of-the-envelope calculations. These approaches are intended to weed-out a subset of the design space that is likely to be dominated by other approaches. Much of this process is accomplished, effectively, through good systems engineering such as the IPPD process [202] and Morphological Analyses. [80]

Once a manageable subset of approach alternatives has been established, however, some form of ranking for those alternatives is usually required while still in the pre-conceptual design phase. In the case of the JCIDS process, the ICD requires a ranked list of approaches prior to the AoA. Because of the complexity of the multiple objectives and attributes associated with the system approaches being ranked, in all likelihood some sort of quantification of measures of effectiveness will have been completed. If a relation between these quantifications and different scenarios can be completed, Global Regret Analysis can be used in pre-conceptual design. Based on the timeline of typical JCIDs studies, which is between six months and a year, this exploration is not considered unreasonable.

Because Global Regret Analysis only is applicable once certain conditions are met in pre-conceptual design, the method will be considered moderately applicable.

6.1.4 Robustness Evaluation at the Capability Level

Global Regret Analysis does a good job of capturing the difference in robustness among candidate alternatives at the capability level. By leveraging the work of Biltgen [24] and Ender [78], [152] for the construction of the analysis framework, the focus for design at all levels of the SoS hierarchy is the military capability. The methodology presented in this dissertation encourages the definition of the local regret fitness metric in terms of the military capability and costs of each SoS alternative. In this way the Global Regret for each SoS alternative is a function of the probability weighted military capability across all scenarios. By casting Global Regret in those terms, the capability level remains most important in the design process.

6.1.5 Applicable to Systems-of-Systems

The applicability of Global Regret Analysis to SoS problems was tackled by considering a SoS problem for the application of the method. The primary differences between the handling of a SoS and a systems problem arise in the modeling and simulation aspects of the design. In particular, the modeling approach chosen for a SoS will tend to favor techniques that focus on the interaction among systems in the hopes of identifying emergent behavior. These models tend to involve the specification of environmental parameters, which can naturally be used as the foundation of the scenario space. In either case, however, because Global Regret Analysis can use measures of merit from any level of the SoS hierarchy, it is applicable to SoS, systems, and subsystems.

6.1.6 Applicable to Multi-Objective Problems

There are two approaches in the current state-of-the-art for multi-objective problems. The first approach uses some form of overall evaluation criterion (OEC), which combines the different objective of the problem into a single score. The challenge associated with creating a valid and effective OEC is assigning appropriate weighting factors to the different dimensions of the problem. The second approach is the use of Pareto Frontiers. Pareto Frontiers carry forward a family of solutions instead of a single "best" solution. These solutions define the hyper-space boundaries of performance with respect to the competing objectives of the design problem [179]. A three-dimensional example of a Pareto Frontier is shown in Figure 132. In the example, each of the three attributes improves by increasing. The family of solutions that represents the boundary of the performance with respect to the three attributes is shown by the convex red surface in the figure.

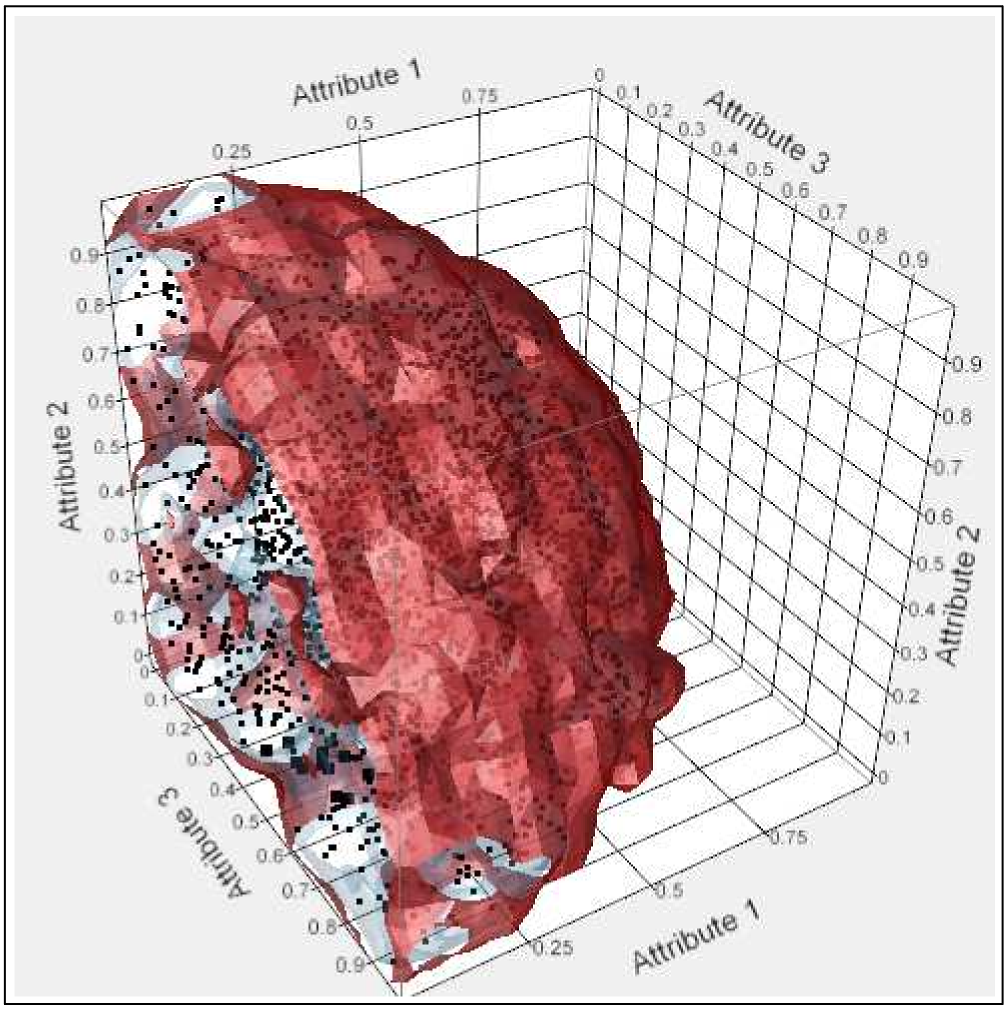


Figure 132: 3-D Pareto Frontier

Global Regret Analysis requires the use of an OEC because of the need to create a single value for comparison with other solutions for a particular scenario. Because only one of the two state-of-the-art approaches to multi-objective problems can be handled in the current formulation, the method will be considered to be mostly applicable to multi-objective problems.

6.1.7 Optimizable

The ability of Global Regret Analysis to be coupled with an optimizer depends on the behavior of the Global and Local Regret functions with changing scenario parameters. As a general rule, the capability of the alternatives under consideration was observed to be non-linear with local minima, and discrete jumps in some regions of the scenario space. This behavior was what drove the use of Artificial Neural Network-based surrogate models. However, because Global Regret Analysis does lead to a single-value measure for each candidate alternative, the analysis can be written in a standard optimization form. [243]

Because of the non-linear behavior of the Global Regret Function, a stochastic optimizer would be the preferred choice, as opposed to a pure gradient-based or path-building method. Both genetic-algorithm [243] and simulated-annealing [75] approaches could be applied to this type of problem, especially given the rapid-response of the surrogate models.

Because fairly advanced stochastic optimization techniques are required for use with most Global Regret Analysis applications, the method will be considered to have mostly met the optimization criteria.

6.1.8 Automated

The applications used in the testing of the hypotheses for this dissertation, and for the complete methodology demonstration have both been conducted in an automated fashion. While there was considerable time required for the construction of the parametric scenario generator for each case, once the environment was built and the alternatives modeled, the DoE was run through an automated process using Model Center. The computational time for the hypothesis demonstration was approximately one week on a desktop computer, which would have been significantly more if the cases had been run manually.

In a sense, Global Regret Analysis requires an automated approach to modeling and simulation because of the large number of scenarios that must be explored to create the scenario surrogate models for each alternative. The degree of automation will impact the time required to complete the analyses, but nothing about the method inhibits automation. There is no human-in-the-loop requirement once the cases have been programmed. The determination of an appropriate objective function does require the interaction with decision makers, and often iteration among interested parties. Therefore, Global Regret Analysis will be considered to be a moderately automated approach.

6.1.9 Applicable to Revolutionary Concepts

Global Regret Analysis is applicable to revolutionary concepts to the degree that they can be captured by the modeling and simulation environment chosen by the designers. Revolutionary concepts imply that simple empirical relationships that have been traditionally used for aircraft conceptual and pre-conceptual design will no longer be applicable. In this case, the designer must rely on physics-based analysis for understanding the capability of the revolutionary concept.

The example used for hypothesis testing in this dissertation used a physics-based modeling and simulation approach. The agents in the environment were modeled based on physical equations derived from basic physical principles such as the Newton's Second Law for calculating aircraft flight performance. Because these relationships do not change for revolutionary concepts, the revolutionary concepts can be integrated easily into the type of modeling and simulation used with Global Regret Analysis.

6.1.10 Robustness Evaluation based on Full Life Cycle

Because of the scope of this dissertation effort, the applicability of Global Regret Analysis to the full life cycle of a System of Systems was not considered in either the example used for hypothesis testing nor the example used for demonstration of the entire method. The lack of an example to cite for determining the applicability of Global Regret Analysis to the full life cycle means that the determination must be made on the basis of logic, and could be disproved in the future.

There are two primary requirements for a designer to use Global Regret Analysis: a single OEC that exists at the capability level (and potentially other levels if requirements dictate) and the ability to consider the impact of different possible scenarios on that OEC. Therefore, if other aspects of the SoS's life cycle, such as maintenance, logistics, training, disposal, etc, can be modeled so that competing alternatives receive a numerical score in each area, and that score can change according to different possible scenarios, Global Regret Analysis will be applicable for assessing the robustness of the entire life cycle of the SoS. However, if models of those life cycle processes can not return a score for each alternative, or the models are not able to incorporate scenario parameters into that score, Global Regret Analysis would not be suited for assessing the robustness of that particular aspect of the SoS life cycle.

Because of the varying level of maturity of models for the entire SoS life cycle, Global Regret Analysis will be considered to have a low-to-moderate ability to capture the robustness of these activities. Additionally, the planning processes for post-acquisition activities of the SoS life cycle are typically not handled until after the Milestone A review in the defense acquisition process. However, as models for the entire SoS life cycle mature and are integrated more readily into conceptual and pre-conceptual design, Global Regret Analysis will be able to incorporate those models without modification to the method.

6.2 Comparison of Global Regret Analysis to Existing Methods

Having discussed the merits of Global Regret Analysis with respect to each of the metrics used for qualitative assessment of robustness methods, Global Regret Analysis was assigned “poor”, “moderate”, or “good” score for each. These scores are summarized in the final column of Figure 133. As can be seen in the figure, for some metrics, such as “Robustness Evaluation at Capability Level,” Global Regret Analysis is clearly the superior option. With respect to some other metrics, such as “Robustness Based on Full Life Cycle”, Global Regret Analysis falls short of other approaches.

Poor - ●
Moderate - ⊙
Good - ○

	Not Optimizer Based			Optimization		
	Kazmer and Roser (1999)	Ford and Barkan's Robust Concept Design (1995)	Taguchi's Parameter Design	Lewis and Parkinson (1994)	Wilde (1992)	Global Regret Analysis
Mathematical Definition of Robustness	○	○	⊙	⊙	○	○
Applicability in Conceptual Design	⊙	○	⊙	●	⊙	○
Applicability in Pre-Conceptual Design	●	⊙	●	●	●	⊙
Robustness Evaluation at Capability Level	●	●	●	●	●	○
Applicable to Systems-of-Systems	⊙	⊙	⊙	⊙	●	○
Applicable to Multi-Objective Problems	○	⊙	⊙	⊙	●	⊙
Optimizable	○	⊙	⊙	○	○	⊙
Automated	●	●	●	○	○	⊙
Applicability to Revolutionary Concepts	●	⊙	●	●	●	○
Robustness Evaluation Based on Full Life Cycle	⊙	○	●	⊙	●	⊙

Figure 133: Comparison of Robustness Methods

In order to gain more insight into the “goodness” of Global Regret Analysis, the Technique for Ordered Preference by Similarity to an Ideal Solution (TOPSIS) methodology was used. TOPSIS is discussed in detail in Appendix A. Two approaches were used to quantifying the data in Figure 133. The first approach assigned a value of 1 to poor performance, 2 to moderate performance, and 3 to good performance. This type of scoring is referred to as a linear scale, and can be useful for its simplicity. The second approach used a ratio scale, where a value of 1 was assigned to poor performance, 3 to moderate performance, and 9 to good performance. This type of scale allows approaches with strengths in a particular area to be easily distinguishable from those with low performance.

Because it is impossible to know which of the metrics used to judge robustness methodologies will be most important to analysts, a large number of weighting schemes were developed using a MCS approach, and the statistics for the TOPSIS "best" solution recorded. The results of the MCS, for which 10,000 cases were run, are shown in Figure 134 and Figure 135 for the linear scale and ratio scale, respectively.

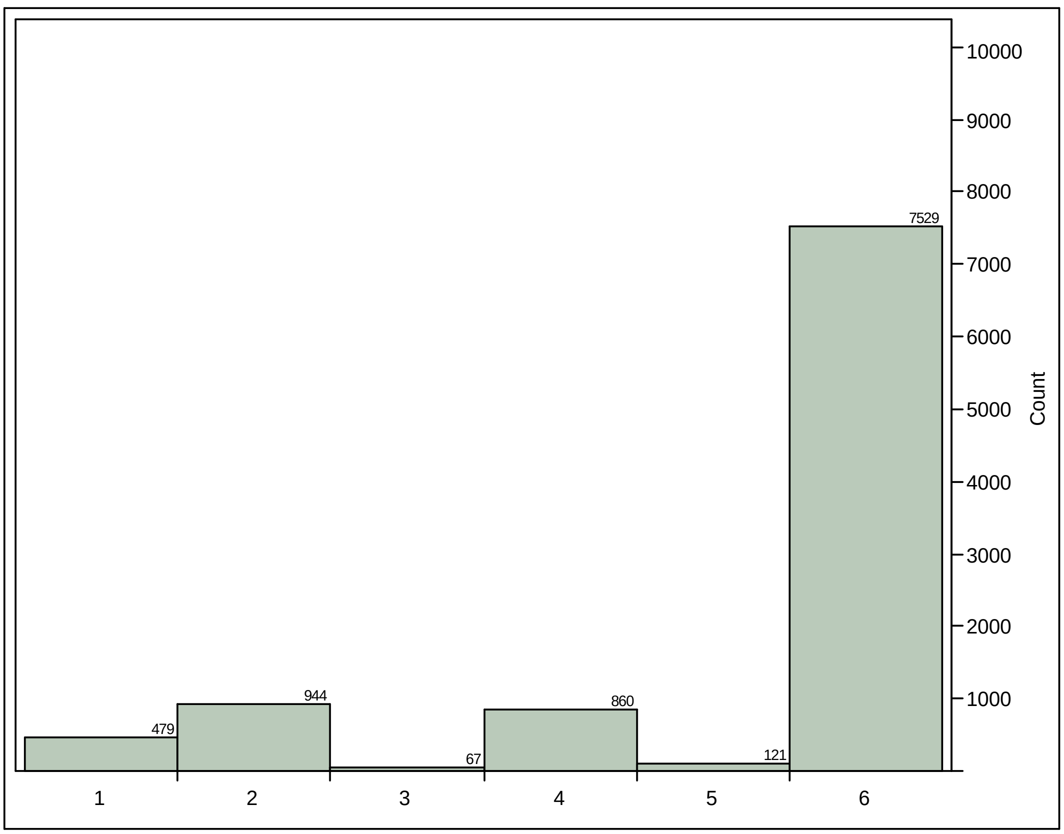


Figure 134: Best Approach, Linear Scale

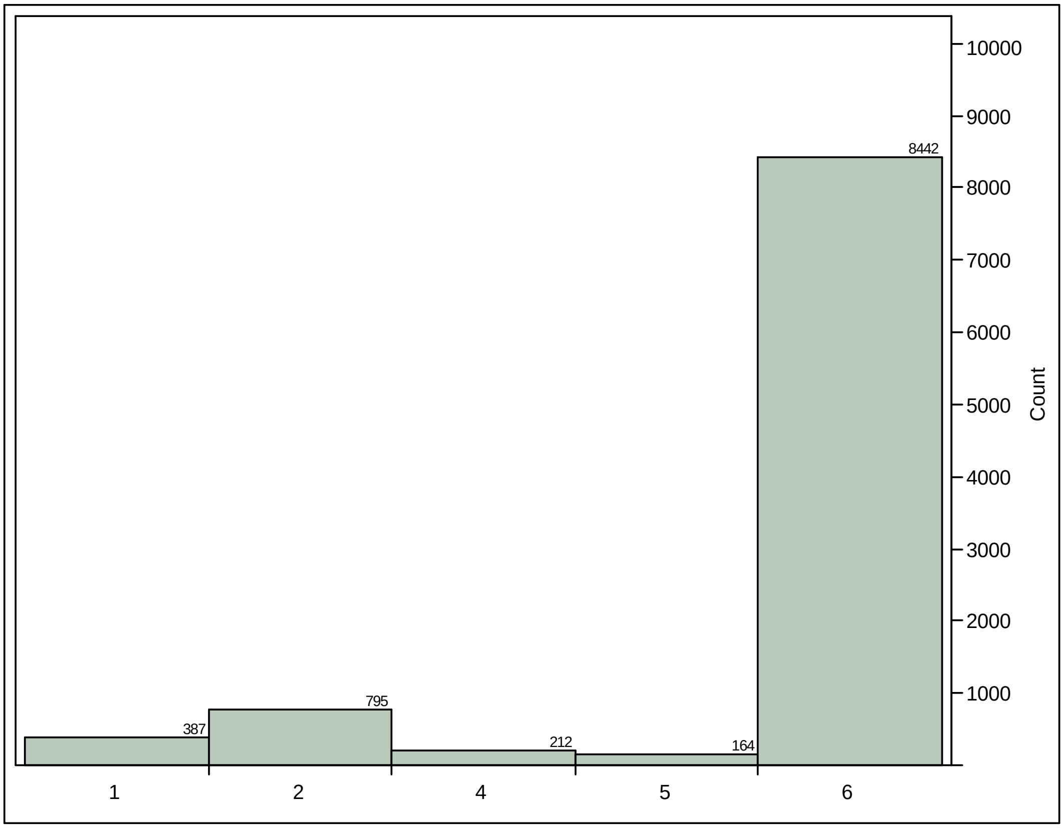


Figure 135: Best Approach, Ratio Scale

As can be seen in the figures, the number of times a particular approach appears as the dominant approach changes depending on the weighting scheme. For example, method 3 does ever appear as a best solution with the ratio scale, though it does with the linear scale. Additionally, the rankings of approaches one through five changes depending on the scale used. What does not change between the scales, however, is the dominance of weightings for which approach 6, Global Regret Analysis, is the “best” approach. While it is clearly not a silver bullet for addressing robustness, the method shows a great improvement over the state-of-the-art for most weighting schemes.

6.3 Contributions

This section contains a summary of the major contributions of this dissertation to the aerospace engineering and defense acquisition fields.

- Survey of the current state-of-the-art with respect to robustness evaluation in defense acquisition
- Identification of gaps in the current defense acquisition process related to robustness evaluation
- Proposal of a method that defines robustness in terms of capability level metrics
- Quantitative definition of robustness that incorporates system and scenario variables
- Supported the individual elements of the method through hypothesis testing
    - Parametric Scenario Generator M&S
    - Surrogate Models for Scenario Space Modeling
    - Mathematical Definition of Global Regret
    - Filtered Monte Carlo Simulation for Visualization
- Demonstrated the use of Global Regret Analysis for the USAF's Persistent Precision Strike application
    - Discovered visualization for single and dual dominance of concepts in local regret visualization
    - Provided a method for visualizing the useful scenario space for different concepts and discussed the implications for mission planning
- Compared Global Regret Analysis to current, state-of-the-art robustness evaluation methods and discussed advantages

6.3.1 INCOSE SoS Challenges

The INCOSE identifies 7 challenges related to the engineering of SoS in the INCOSE Systems Engineering Handbook [123]. These challenges are listed as being above and beyond the challenges associated with engineering conventional systems. Biltgen summarizes the challenges and discusses their contribution to the unique nature of SoS [24].

1. System Elements Operate Independently
2. Systems Elements Have Different Life Cycles
3. The Initial Requirement are Likely to be Ambiguous
4. Complexity is a Major Issue
5. Management Can Overshadow Engineering
6. Fuzzy Boundaries Cause Confusion
7. SoS Engineering is Never Finished

Because the defense acquisition process was the starting point for the development of the ideas of this dissertation, these 7 challenges were not included in the initial development of Global Regret Analysis. However, Global Regret Analysis does partially address two of the challenges identified by the INCOSE: ambiguity of initial requirements and management overshadowing engineering.

Ambiguity in the initial set of requirements can be partially addressed by Global Regret Analysis by structuring the Parametric Scenario Generator and M&S so that the requirements with ambiguity are included as scenario variables. For example, if the radar signature requirements for an aircraft were ambiguous early in the design phase, they could be incorporated into the Global Regret calculation in one of two ways. The first way would be the construction of the local fitness function in such a fashion as to impose a penalty when the signature constraint was violated. This constraint could then be varied and the robustness of the concepts would include the moving requirement. A second approach would be to re-size the set of concepts each time a vehicle requirement changed. This approach would allow the comparison of the robustness of approaches, not just individual concepts. Unfortunately, this approach would be much more computationally intensive than the modified fitness function.

The second challenge that Global Regret Analysis partially addresses is the challenge of management overshadowing engineering. As was shown in the demonstration of the methodology on the Persistent Precision Strike example, for each of the five concepts a combination of OEC weighting factors and scenario variables could be found to make any concept dominant. The implication of this discovery is that if program management had a concept that was preferred for reasons beyond those stated for the analysis, a case could be made for that concept being the "best." However, by using the Global Regret Analysis approach in assessing robustness, those few cases could quickly be shown as outliers, while the dominant concepts were "best" over the majority of the scenario space. While Global Regret Analysis does address the issue of management being able to dial in scenarios for a favored candidate, it does not help with another aspect of the management issue. Because SoS are so complex, the program management necessary to coordinate among the different vendors, designers, users, and funders can be overwhelming.

6.4 Final Thoughts

Finding robust solutions to problems requires much more effort than finding an optimal solution. The optimal solution makes the designer only think of one scenario, one (sometimes composite) objective function, and a set of constraints. The robust solution, on the other hand, needs to consider the objective function, the set of constraints, and how those change over many possible scenarios. But, given the rapid pace at which the world is changing and the lives that are on the line when considering military systems, we must strive to overcome these difficulties and make the best possible decisions.

6.4.1 Future Work

The majority of the work in this dissertation focused on the comparison of a few, materiel solutions to a capability need. However, the process of parsing through the billions of morphological combinations that exist in the design space, especially when non-materiel solutions are including, has not been well documented in the defense acquisition process. A formalization of the morphological analysis methods for the early defense acquisition process could greatly clarify the process of reducing the design space in a logical and systematic way.

There is the potential to greatly improve the effectiveness of a system-of-systems through the tactical manner in which it is employed. US Special Forces use essentially the same equipment as the enemies they engage, yet are so tactically superior they nearly always prevail. The tactics for using a system are typically left to they operational phase of the system's life cycle. Beyond a brief CONOPS, the soldiers must train and learn to employ the system after its design. If tactics could be accounted for easily in the early modeling and simulation efforts of conceptual design, those tactics could in turn be accounted for in the design itself, possibly reducing cost on increasing expected performance.

The Global Regret Analysis methodology could be expanded and tested in a number of ways that would contribute to its usefulness. The first area for additional research is the implementation of the fitness function for the evaluation of the local regret. While OECs have wide acceptance in the aerospace and defense acquisition communities, they are not without critics. Exploration of the use of other techniques such as AHP, TOPSIS, or Pareto Optimality would enhance the ability of Global Regret Analysis to overcome criticisms of OECs.

The second area of additional research for the Global Regret Analysis methodology would be improvements in the definition for the local likelihood function for the scenario space. In the formulation presented for the local likelihoods, all of the scenario variables are treated as independent variables. In reality, however, there would most likely be correlations among the various scenario variables. In many cases these correlations could be linked to other metrics, such as levels of funding in the adversary's research and development, level of hostility, and environmental characteristics of the adversary's country. Accounting for these correlations would allow for a more accurate vision of the likelihood of plausible futures, and consequently, a better assessment of the Global Regret for each competing alternative.

6.4.2 Support for Global Regret Analysis

Throughout this dissertation, the author has strove to address the creation of the Global Regret Analysis method in a scientifically sound, repeatable way. However, especially with regard to the experiments in logic that led to the selection of the particular tools used, the human mind was the laboratory for the experiment. In these cases, other scientists and engineers might reach different conclusions. Therefore, for consensus to be reached, scientific dialog on the topic is required. Especially when the hypotheses of the work deal with subjects that can not be explicitly proven, the merits and weaknesses of the method must be explored.

This dissertation provided two example problems that provided support to the overall hypothesis that Global Regret Analysis has strength for addressing robustness early in the defense acquisition process. The examples were selected to hopefully allow readers to extrapolate the metrics, systems, and environments considered to problems relevant to their own work. However, because the degree pursued by the author is in aerospace engineering, it was necessary to keep the example applications in that arena. The work contained here has provided two successful applications of Global Regret Analysis; however, the bounds of its applicability still need to be tested. The bound of Global Regret Analysis's applicability have been suggested based on logic throughout the development of the method. However, these bounds will only become clear as the method is used.

6.4.3 Exploring the Bounds of SoS Performance

A wide variety of factors beyond simply the SoS's capability and cost drive defense acquisitions. The military industrial complex involves lobbyists, politicians, military planner, and budget controllers. [68] In most cases, elected officials must show how the work they have done has benefited their home district, adding another degree of complexity to the SoS acquisition problem. While Global Regret Analysis does a good job of allowing decision making based on the capability and cost concerns of the SoS, M&S does not exist for all of the political factors of the military industrial complex.

Regardless of these complexities, Global Regret Analysis provides one major benefit for decision makers, even if decisions are not truly based on the capability and cost of the SoS. The ability to understand the bounds of performance for a system relative to its alternatives would be very valuable for a military campaign. If mission planners understood the topology of the scenario space and were able to explore that space rapidly, situations that posed challenges could be addressed before lives were lost. If after the acquisition process was complete, it became apparent that the SoS would be used in an area of degraded performance as identified by Global Regret Analysis, training or tactics could be altered to improve the system performance, or the mission might be cancelled. This understanding would also help politicians prepare for potential backlash by identifying a priori the military situations that might result in regret.

APPENDIX A

METHODS

A.1 Functional Decomposition

A functional decomposition, just as the name suggests, breaks down the system design based on the functions it will accomplish. Rather than looking at the physical system parts, the functional decomposition could be accomplished by asking "what must the system do?" An example functional decomposition for an aircraft might include "generate lift", "store payload", and "generate thrust." Functional decompositions are useful for organizing conceptual design alternatives when the requirements do not dictate a specific type of physical system. They encourage engineers to "think outside the box" when deciding how to accomplish the tasks required of the system-of-systems.

A.2 DoE Methods

Design of Experiments is a collection of mathematical approaches to structuring experiments in such a way as to gain as much information from each experimental run as possible. According to Breyfogle, "DoE Techniques offer a structured approach to change many factor settings within a process at once and observe the data collectively for improvements/degradations." [36] There are many types of DoEs that are specialized for certain applications, so some knowledge of the nature of the system being explored is helpful. When little knowledge of the system is available, a set of screening DoEs can be used to gain insight with little effort. DoEs are most helpful when individual experimental runs are expensive, either computationally, in terms of labor, or dollars.

Reductionism is often coupled with DoE techniques. In SoS problems especially, the number of independent variables available for manipulation is large. Screening DoEs, coupled with Analysis of Variance (ANOVA), can help reduce the number of variables that are considered in early design phases. [250] In most SoS problems, only a handful of variables have significant influence on the variability of the metrics of interest over the ranges relevant to the problem. By using ANOVA, those variables may be identified and the rest defaulted with little impact on the accuracy surrogate models.

A.3 TOPSIS

The Technique for Ordered Preference by Similarity to an Ideal Solution (TOPSIS) provides a simple procedure for obtaining a definitive set of ranked alternatives. TOPSIS involves the selection of important design criteria, the assignment of weights to those criteria (often from the QFD) and then the evaluation of alternatives based on the distance of those alternatives' design criteria from an ideal. For this procedure a positive and negative "ideal" case are created, and the alternative that is closest to the positive ideal and farthest from the negative ideal is ranked the highest. Geometrically this is shown in Figure 136. The Pareto Curve is a curve defined by the limit of cases that minimize attributes 1 and 2. The Pareto Curve can be thought of as an isovalue contour of the maximum fitness for the feasible design space based on the overall evaluation function that is used to evaluate concepts. The best case would be that case which minimized attributes 1 and 2 based on their respective weights, and lies as close to the ideal solution as possible.

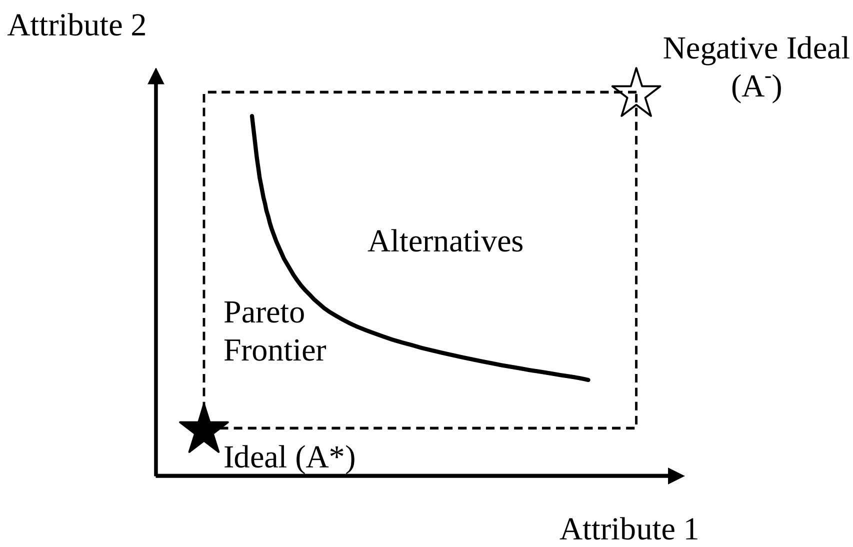


Figure 136: Graphical Representation of TOPSIS

TOPSIS involves first creating a decision matrix of alternatives as rows and responses/characteristics as columns. These columns are then populated for each alternative based either on simulation results, empirical data, or qualitative assessment. If a qualitative assessment is used, the values must be converted to a numerical scale. The dimensional values in this matrix are then normalized on a scale of -1 to 1 by dividing each entry by the square-root of the sum-of-squares of its column and then multiplied by the weighting factor associated with the response or characteristic. Each response is then characterized as a cost or a benefit, for example payload and endurance are benefits and are maximized while carbon dioxide emissions and landing field length are costs and are minimized. The responses considered for this implementation follow a weighting structure that is adjustable depending on the judgment of the evaluating engineers.

In order to establish a basis for comparison for each alternative a positive and negative ideal are selected based on the weighted values of the responses. The positive ideal is the set that includes the maximum value in the matrix for each of the benefits and the minimum value for each of the costs. Then negative ideal possesses the maximum value for each of the costs and the minimum value for each of the benefits. The separation of each alternative from the positive and negative ideal is calculated based on the square-root of the sum-of-squares of the differences between the alternative and the positive ideal, and then again for the alternative and the negative ideal.

The relative closeness of each alternative to the ideal solution is then calculated based on the separation from the negative solution divided by the sum of the positive and negative separations. This ranking system results in a value for each alterative between 0 and 1, where 0 corresponds to the worst alternative and 1 to the best. A flow chart of the TOPSIS methodology is included as Figure 137.

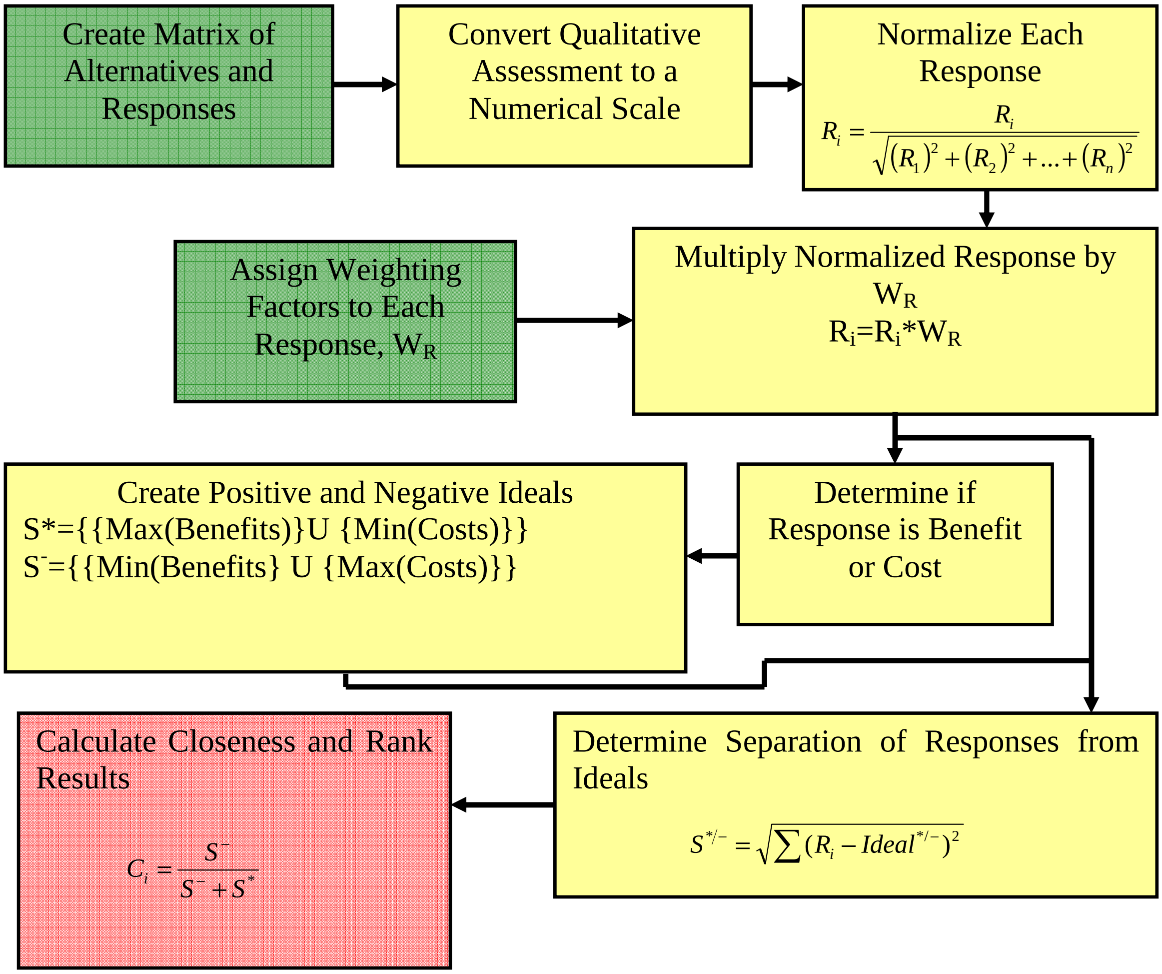


Figure 137: TOPSIS Flowchart

When the fitness of each concept with respect to the system metrics used for concept evaluation are plotted on each of the axes of a radar plot, the best system from the TOPSIS methodology will correspond roughly to that which has the most area in the radar plot. Because of the dependence of the overall evaluation criterion on the weights assigned to each system evaluation metric, it is important to have consensus on the weights from the evaluating engineers, or create a probabilistic assessment by varying the weights over a large number of cases and observing trends in the resulting rankings. This type of analysis can be used to show the sensitivity of system ranking to requirement weights and allows the further study of significant requirements. An example implementation of a TOPSIS evaluation for the selection of a UAV system is shown in Figure 138.

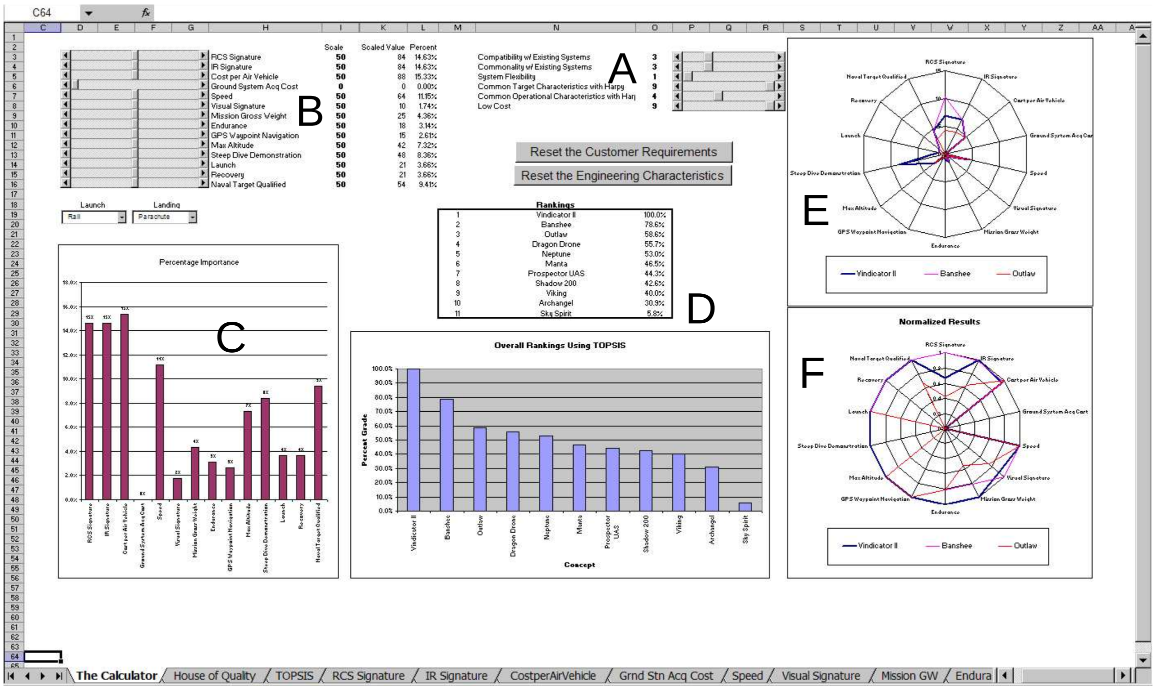


Figure 138: TOPSIS Decision Making Tool

APPENDIX B

CODES

B.1 MATLAB Terrain Generators

B.1.1 Ridgemaker

```
%Ridge Generator - Fractal
%Benjamin Poole

function [terrain45] = ridgemaker(mountainousness,roughness,max_elev)

% mountainousness = 1;
% roughness = .3;

InitialLine = [0,0;1,0];
Line = InitialLine;
n=1;
randomrange = mountainousness;
while n <= 7
   segments = size(Line);
   segments = segments(1)-1;
   segcounter = 1;
   while segcounter <= segments
      distance = (Line(segcounter+1,1) - Line(segcounter,1));
      midpoint = 0.5*(Line(segcounter+1,1) - Line(segcounter,1))+ Line(segcounter,1);
      slope = (Line(segcounter+1,2) - Line(segcounter,2))/distance;
      mid_elev = slope*(midpoint-Line(segcounter,1))+ Line(segcounter,2);
      displacement = rand*randomrange;
      direction = 1;
      if rand<0.5
         direction = -1;
      end
      newpoints(segcounter,:) = [midpoint, direction*displacement+mid_elev];
      segcounter = segcounter + 1;
   end
   Line = [Line;newpoints];
   Line = sortrows(Line);
   randomrange = randomrange*roughness;
   n = n+1;
end
% plot(Line(:,1),Line(:,2))
% AXIS([InitialLine(1,1) InitialLine(2,1) -1 1])
terrain1 = Line(:,2);
```

```
terrain2 = max_elev*terrain1;
terrain45 = -1*min(terrain2)+ terrain2;
```

B.1.2 Terrainmaker

```
%Terrain Maker
%Benjamin Poole


function [Terrain]=terrainmaker1(mountainousness,roughness,downpercent,iterations)

%clear all
close all
clc


%mountainousness = 1;
%roughness = 0.5;
Initial_Square = [0,0,0;0,1,0;1,0,0;1,1,0]; %x,y,z
%downpercent =.3;

%initializations
Terrain = Initial_Square;
randomrange = mountainousness;
number_squares = 1;
%iterations =5;

%Uncomment if you want the same terrain repeatedly
%rand('state',0)

iteration_counter = 0;
while iteration_counter<iterations
   %Square Step
   n = 0;
   while n<number_squares
      %Find the cornerpoints of the square for this sub-iteration
      gridsize = (number_squares^.5);%*(Initial_Square(3,1)-Initial_Square(1,1))
      Square1 = [mod(n,gridsize)/gridsize*(Initial_Square(3,1)-
Initial_Square(1,1)),(floor(n/gridsize))/gridsize*(Initial_Square(2,2)-
Initial_Square(1,2))];
      Square1 = Terrain(find(Terrain(:,1)==Square1(1)&Terrain(:,2)==Square1(2)),:);
      Square2 = [Square1(1)+1/gridsize,Square1(2)];
      Square2 = Terrain(find(Terrain(:,1)==Square2(1)&Terrain(:,2)==Square2(2)),:);
      Square3 = [Square1(1),Square1(2)+1/gridsize];
      Square3 = Terrain(find(Terrain(:,1)==Square3(1)&Terrain(:,2)==Square3(2)),:);
      Square4 = [Square1(1)+1/gridsize,Square1(2)+1/gridsize];
```

```
        Square4 = Terrain(find(Terrain(:,1)==Square4(1)&Terrain(:,2)==Square4(2)),:);
        Square = [Square1;Square2;Square3;Square4];

        %Find the Midpoint Height
        average_elev = (Square(1,3)+Square(2,3)+Square(3,3)+Square(4,3))/4;
        displacement = rand*randomrange;
        direction = 1;
        if rand<downpercent
           direction = -1;
        end
        midpoint1 = (Square(2,1)-Square(1,1))*0.5+Square(1,1);
        midpoint2 = (Square(3,2)-Square(2,2))*0.5+Square(1,2);
        midpoint3 = direction*displacement+average_elev;
        midpoint = [midpoint1,midpoint2,midpoint3];
        midpoints(n+1,:) = [midpoint];
        %Update the Terrain
        Terrain = [Terrain;midpoint];
        n=n+1;
    end
    %diamond step
    m=0;
    while m<number_squares
        d=0;
       while d<4
           %find corner points of the diamond for this sub-iteration
           if d==0
               Diamond1 = [midpoints(m+1,1),midpoints(m+1,2)-
(1/gridsize)*(Initial_Square(2,2)-Initial_Square(1,2))];
               Diamond2 = [midpoints(m+1,1)-(1/gridsize/2)*(Initial_Square(3,1)-
Initial_Square(1,1)),midpoints(m+1,2)-(1/gridsize/2)*(Initial_Square(2,2)-
Initial_Square(1,2))];
               Diamond3 = [midpoints(m+1,1), midpoints(m+1,2)];
               Diamond4 = [midpoints(m+1,1)+(1/gridsize/2)*(Initial_Square(3,1)-
Initial_Square(1,1)),midpoints(m+1,2)-(1/gridsize/2)*(Initial_Square(2,2)-
Initial_Square(1,2))];
           end
           if d==1
               Diamond1 = [midpoints(m+1,1)-(1/gridsize/2)*(Initial_Square(3,1)-
Initial_Square(1,1)),midpoints(m+1,2)-(1/gridsize/2)*(Initial_Square(2,2)-
Initial_Square(1,2))];
               Diamond2 = [midpoints(m+1,1)-(1/gridsize)*(Initial_Square(3,1)-
Initial_Square(1,1)),midpoints(m+1,2)];
               Diamond3 = [midpoints(m+1,1)-(1/gridsize/2)*(Initial_Square(3,1)-
Initial_Square(1,1)),midpoints(m+1,2)+(1/gridsize/2)*(Initial_Square(2,2)-
Initial_Square(1,2))];
               Diamond4 = [midpoints(m+1,1), midpoints(m+1,2)];
```

```
        end
        if d==2
            Diamond1 = [midpoints(m+1,1), midpoints(m+1,2)];
            Diamond2 = [midpoints(m+1,1)-(1/gridsize/2)*(Initial_Square(3,1)-
Initial_Square(1,1)),midpoints(m+1,2)+(1/gridsize/2)*(Initial_Square(2,2)-
Initial_Square(1,2))];
            Diamond3 =
[midpoints(m+1,1),midpoints(m+1,2)+(1/gridsize)*(Initial_Square(2,2)-
Initial_Square(1,2))];
            Diamond4 = [midpoints(m+1,1)+(1/gridsize/2)*(Initial_Square(3,1)-
Initial_Square(1,1)),midpoints(m+1,2)+(1/gridsize/2)*(Initial_Square(2,2)-
Initial_Square(1,2))];
        end
        if d==3
            Diamond1 = [midpoints(m+1,1)+(1/gridsize/2)*(Initial_Square(3,1)-
Initial_Square(1,1)),midpoints(m+1,2)-(1/gridsize/2)*(Initial_Square(2,2)-
Initial_Square(1,2))];
            Diamond2 = [midpoints(m+1,1), midpoints(m+1,2)];
            Diamond3 = [midpoints(m+1,1)+(1/gridsize/2)*(Initial_Square(3,1)-
Initial_Square(1,1)),midpoints(m+1,2)+(1/gridsize/2)*(Initial_Square(2,2)-
Initial_Square(1,2))];
            Diamond4 = [midpoints(m+1,1)+(1/gridsize)*(Initial_Square(3,1)-
Initial_Square(1,1)),midpoints(m+1,2)];
        end
        %does this diamond already exist? (check to see if midpoint is in
        %terrain)
        midpoint_d = Diamond1+(Diamond3-Diamond1)/2;
        exist_test = find(Terrain(:,1)==midpoint_d(1)&Terrain(:,2)==midpoint_d(2));
        if isempty(exist_test)
          Diamond1 =
Terrain(find(Terrain(:,1)==Diamond1(1)&Terrain(:,2)==Diamond1(2)),:);
          Diamond2 =
Terrain(find(Terrain(:,1)==Diamond2(1)&Terrain(:,2)==Diamond2(2)),:);
          Diamond3 =
Terrain(find(Terrain(:,1)==Diamond3(1)&Terrain(:,2)==Diamond3(2)),:);
          Diamond4 =
Terrain(find(Terrain(:,1)==Diamond4(1)&Terrain(:,2)==Diamond4(2)),:);
          Diamond = [Diamond1;Diamond2;Diamond3;Diamond4];
          direction = 1;
          if rand<downpercent
              direction = -1;
          end
          displacement = rand*randomrange;
          midpoint_d = [midpoint_d,direction*displacement+mean(Diamond(:,3))];
          Terrain = [Terrain;midpoint_d];
        end
```

```
            d = d+1;
        end
        m=m+1;
    end
    number_squares = number_squares*4;
    iteration_counter = iteration_counter+1;
    randomrange = randomrange*roughness;
end
Terrain;

%plotter
x_counter = 0;
x=[];
while x_counter<=(2*gridsize)
    x                    =                    [x;Initial_Square(1,1)+x_counter*(Initial_Square(3,1)-
Initial_Square(1,1))/(2*gridsize)];
    x_counter = x_counter+1;
end
y_counter = 0;
y=[];
while y_counter<=(2*gridsize)
    y                    =                    [y;Initial_Square(1,2)+y_counter*(Initial_Square(2,2)-
Initial_Square(1,2))/(2*gridsize)];
    y_counter = y_counter+1;
end
Z = [];
for j=1:(2*gridsize+1)
    for i=1:(2*gridsize+1)
        Z(i,j) = Terrain(find(Terrain(:,1)==x(j)&Terrain(:,2)==y(i)),3);
    end
end
 surf(x,y,Z)
 hold;
```

B.2 MATLAB LOS Calculators

B.2.1 Urban LOS

```
%urban coverage 3-D
%Benjamin Poole


%ALL DISTANCES IN METERS

clear all
close all
clc

%aircraft altitude, sensor width
h=100000;
w=1000;

%average building size, height, street width
a=10;
b=10;
c=8;

%worst case viewing (over center of the block)
v=a/2;

%number of blocks in swath
blocks=w/(a+b);

%counter initialization
counter = 0;
covered_area=0;

while counter<blocks/2
   x=v+b;
   s=b*x-1/2*v*c/(h-c)*x*(1+x^2/v^2)^(1/2)-1/2*v*c/(h-
c)*log(1/v^2*x/(1/v^2)^(1/2)+(1+x^2/v^2)^(1/2))/(1/v^2)^(1/2);
   s=s*4;
   x=v;
   s1=b*x-1/2*v*c/(h-c)*x*(1+x^2/v^2)^(1/2)-1/2*v*c/(h-
c)*log(1/v^2*x/(1/v^2)^(1/2)+(1+x^2/v^2)^(1/2))/(1/v^2)^(1/2);
   s1=s1*4;
   if s<0
      s=0;
      if s1<0
         s1=0;
```

```
        end
    end
    avgdepth=s1/4/v;
    slightly_optomistic_sections=(counter)*8*avgdepth*a;
    covered_area = covered_area+s+s1+slightly_optomistic_sections;
    v=v+a+b;
    counter=counter+1;
end
covered_area=covered_area
covered_fraction=covered_area/w^2
covered_percent=covered_fraction*100

covered_of_available = covered_area/(((1-(a/(a+b))^2))*w^2)
covered_of_available_percent = covered_of_available*100
```

B.2.2 Terrain LOS

```
%terrain coverage 2-D
%Benjamin Poole
%inputs are h (height of aircraft), halfswath (length of half of the swath
% and terrain, a vector of the terrain heights

function [visibility_amount] = linear_terraincoverage(h,halfswath,terrain)

%h is height agl
% h = 100;
% halfswath = 100;
% terrain = [3,4,50,12,13,14,10,1,2,3,4];

d = halfswath/(length(terrain)-1);

x1=0;
x2=x1+1;

h1 = terrain(x1+1);
h2 = terrain(x2+1);

angle2 = atan(x2*d/(h-h2));
angle1 = atan(x1*d/(h-h1));
theta = angle2 - angle1;

hmax = [h2,d];

if h1 > h2 & theta < 0
    visibility(x2) = 0;
```

```
else
    visibility(x2) = 1;
end

x1=x2;
x2=x1+1;


 while x2 <= halfswath/d
     h1 = terrain(x1+1);
     h2 = terrain(x2+1);
     visibility(x2)=1;

     if h1 < hmax | h2 < hmax
       angle_max = atan(hmax(2)/(h-hmax(1)));
       angle2 = atan(x2*d/(h-h2));
       angle1 = atan(x1*d/(h-h1));
       theta1 = angle1-angle_max;
       theta2 = angle2-angle_max;
       if theta1 < 0 | theta2 <0
           visibility(x2)=0;
       end
     end

     if visibility(x2)==1
         angle2 = atan(x2*d/(h-h2));
         angle1 = atan(x1*d/(h-h1));
         theta = angle2 - angle1;
         if h1 > h2 & theta < 0
             visibility(x2) = 0;
         end
     end

     if h2>hmax(1)
         hmax = [h2,x2*d];
     end
     x1=x2;
     x2=x1+1;
 end
 visibility;
 visibility_amount = sum(visibility)/(length(visibility));
```

B.3 NetLogo Model

```
;Hunter-Killer Model
;By Benjamin Poole
;18 Apr 2008

;===================================================================
;Declarations
;===================================================================

breed [HKs HK]
breed [hunters hunter]
breed [killers killer]
breed [missiles missile]
breed [targets target]

globals [
missile-speed
detonation-radius
max-missile-flight
rearming-counter
refueling-counter
blue-missiles-counter
red-missiles-counter
]

HKs-own [
objective
nearest
speed
decider?
discernability-timer
permission?
missile-inbound?
hiding?
weapons
fuel
resupply-timer
resupplying?
last-x
last-y
last-heading
weapon-range
returning?
fly-back?
last-heading-fb
```

```
]

hunters-own [
objective
nearest
speed
decider?
discernability-timer
permission?
missile-inbound?
hiding?
weapons
fuel
resupply-timer
resupplying?
last-x
last-y
last-heading
nearest-support-dist
nearest-support
returning?
]

killers-own [
objective
nearest
speed
decider?
discernability-timer
permission?
missile-inbound?
hiding?
weapons
fuel
resupply-timer
resupplying?
last-x
last-y
last-heading
weapon-range
]

missiles-own [
objective
speed
flight-time]
```

```
targets-own [
hiding?
hiding-counter
can-hide?
missile-inbound?
variety
loiter-time
popup-percent
discernability   ; value from 1-10 with 10 difficult
armed
sensor-range
weapon-range
objective
speed
permission?
]

patches-own [
urban?]

;================================================================
;Basic codes
;================================================================
to setup
  ca
  random-seed seed
  make-HKs
  make-hunters
  make-killers
  make-targets
  make-patches
  set missile-speed 15
  set detonation-radius 1
  set-default-shape missiles "missile"
  set remember-percentage 0
  set base-decision-time 1
  set max-missile-flight 100
  set refueling-counter 0
  set rearming-counter 0
  set blue-missiles-counter 0
  set red-missiles-counter 0
end

to go
  missile-flight
```

```
  defenses
  huntkill
  target-move
  unhide
  tick
  if (count HKs + count killers) = 0 or count targets with [variety != "SAM"] = 0 or ticks
>= 5000 [
  finish
  stop]
end

;================================================================
;Setup codes
;================================================================

to make-HKs
  set-default-shape HKs "hk"
  create-HKs nHKs [
  set size 10
  set color blue - 2
  set heading 0
  set speed 3
  set decider? autonomous?
  set objective nobody
  set permission? false
  set missile-inbound? false
  set hiding? false
  set weapons max-hk-weapons
  set fuel max-hk-fuel
  set resupply-timer 0
  setxy  (who * 200) / count HKs - 100 + sensor-width / 2 (who * 200) / count HKs - 100
+ sensor-width / 2
  set last-x xcor
  set last-y ycor
  set last-heading heading
  set weapon-range BWRange
  set resupplying? false
  set returning? false
  set fly-back? false]
end

to make-hunters
  set-default-shape hunters "rq-4a"
  create-hunters nHunters [
  set size 10
  set color blue - 2
```

```
  set heading 0
  set speed 3
  set decider? autonomous?
  set objective nobody
  set permission? false
  set missile-inbound? false
  set hiding? false
  set weapons 0
  set fuel max-hunter-fuel
  set resupply-timer 0
  setxy random-xcor random-ycor
  set last-x xcor
  set last-y ycor
  set last-heading heading
  set resupplying? false
  set returning? false]
end

to make-killers
  set-default-shape killers "killer"
  create-killers nKillers [
  set size 4
  set color blue - 2
  set heading 0
  set speed 3
  set decider? autonomous?
  set objective nobody
  set permission? false
  set missile-inbound? false
  set hiding? false
  set weapons max-killer-weapons
  set fuel max-killer-fuel
  set resupply-timer 0
  setxy random-xcor random-ycor
  set last-x xcor
  set last-y ycor
  set last-heading heading
  set weapon-range BWrange
  set resupplying? false]
end

to make-targets
  create-targets ntrucks [
    set variety "truck"
    set shape "truck"
    set hiding-counter 0
```

```
    set size 5
    set hiding? true
    set color gray
    set missile-inbound? false
    set discernability Dtruck
    set popup-percent Ptruck
    set loiter-time Ltruck
    set can-hide? true
    set objective nobody
    set permission? false
    set speed 1
    setxy random-xcor random-ycor]
  create-targets nfacilities [
    set variety "facility"
    set shape "facility"
    set hiding-counter 0
    set size 5
    set hiding? false
    set color red
    set missile-inbound? false
    set discernability 0
    set popup-percent 3
    set loiter-time 15
    set can-hide? false
    set objective nobody
    set permission? false
    setxy random-xcor random-ycor]
  create-targets nSAMs [
    set variety "SAM"
    set shape "sa-6 sam"
    set hiding-counter 0
    set size 5
    set hiding? true
    set color gray
    set missile-inbound? false
    set discernability DSAM
    set popup-percent PSAM
    set loiter-time LSAM
    set can-hide? true
    set armed "SAMissile"
    set sensor-range RSSAM
    set objective nobody
    set permission? false
    set weapon-range RWSAM
    setxy random-xcor random-ycor]
end
```

```
to make-patches
  ask patches [
    ifelse random 100 < urban-percent [
      set urban? true
      set pcolor 5 ]
      [
      set urban? false
      set pcolor 53]]
end

;==============================================================
;HK-killer codes
;==============================================================

;need to work out a way for the permission not to reset if the objective doesn't hide

to huntkill
  ask HKs [
    ifelse weapons = 0 or fuel <= 0 or resupplying? [
      if not resupplying? [
        set last-x xcor
        set last-y ycor
        set last-heading heading]
      resupply] [
    ifelse objective != nobody and decider? [
      fire] [
    ifelse objective != nobody and permission? [

      fire] [
    ifelse objective != nobody and not decider? [
      get-permission] [
    ifelse fly-back? [
      fly-back] [
    hunt]
    ]]]]
    if fuel > 0 [
    set fuel fuel - 1]]

  ask hunters [
    ifelse fuel <= 0 or resupplying? [
      if not resupplying? [
        set last-x xcor
        set last-y ycor
        set last-heading heading]
      resupply] [
```

```
    ifelse objective != nobody and decider? [
      call-in] [
    ifelse objective != nobody and permission? [
      call-in] [
    ifelse objective != nobody and not decider? [
      get-permission] [
    hunt]
  ]]]
  if fuel > 0 [
    set fuel fuel - 1]]

  ask killers [
    ifelse weapons = 0 or fuel <= 0 or resupplying? [
      if not resupplying? [
        set last-x xcor
        set last-y ycor
        set last-heading heading]
      resupply] [
    ifelse objective != nobody and decider? [
      fire] [
    ifelse objective != nobody and permission? [
      fire] [
    ifelse objective != nobody and not decider? [
      get-permission] [
    loiter]]]]
    if fuel > 0 [
    set fuel fuel - 1]]
end

to hunt
    target-available
    ifelse objective != nobody [] [  ; if i have a target, do nothing, else move the aircraft in
one of the ways below
    ifelse returning? [return] [
    ifelse heading = 0 and max-pycor - ycor < speed and max-pxcor - xcor < sensor-width
[
      return] [
    ifelse heading = 180 and ycor = min-pycor and xcor = max-pxcor [
        return] [
    move-aircraft]
      ]]]
end

to return
  ifelse ycor < min-pycor + max list speed sensor-width and xcor < min-pxcor + max list
speed sensor-width [
```

```
  set heading 0
  setxy min-pxcor + sensor-width / 2  min-pycor + sensor-width / 2
  set returning? false][
  set returning? true
  facexy min-pxcor min-pycor
  forward speed]
end

to move-aircraft  ; moves aircraft in the search pattern
  if heading != 180 and heading != 0 [
   set heading last-heading]
  ifelse heading = 0 and max-pycor - ycor >= speed [
   forward speed] [ifelse heading = 0 and max-pycor - ycor < speed  [
    set heading 90
    forward sensor-width
    set heading 180] [ifelse heading = 180 and min-pycor - ycor <= -1 * speed [
      forward speed] [ifelse heading = 180 and min-pycor - ycor > -1 * speed [
      set heading 90
      forward sensor-width
      set heading 0][set heading 0]
      ]]]
end

to target-available ; checks if there is a target within range, if not, sets objective to nobody
 ifelse any? targets with [not hiding? and not member? self turtle-set (list [objective] of hunters)] [
  compute-dist-to-nearest
  let terrain-of-target find-terrain
  ifelse terrain-of-target = true [
   if nearest < sqrt (sensor-width ^ 2 + sensor-length ^ 2) and random 100 < urban-visibility and random 100 < terrain-visibility [
    set objective min-one-of targets with [not hiding? and not member? self turtle-set (list [objective] of hunters)] [distance myself]]] [
   if nearest < sqrt (sensor-width ^ 2 + sensor-length ^ 2) and random 100 < terrain-visibility  [
    set objective min-one-of targets with [not hiding? and not member? self turtle-set (list [objective] of hunters)] [distance myself]]]][
  set objective nobody]
end

to-report find-terrain
 let terrain-of-target [urban?] of patch-set [patch-here] of min-one-of targets with [not hiding? and not member? self turtle-set (list [objective] of hunters)] [distance myself]
 report terrain-of-target
end
```

```
to compute-dist-to-nearest
  let target-of-interest min-one-of targets with [not hiding? and not member? self turtle-set
(list [objective] of hunters)] [distance myself]
  if target-of-interest != nobody [ set nearest distance target-of-interest ]
end

to fire ; if my distance to objective is greater than my weapon range, move toward
objective, else fire (remove, set objective and perm)
  ifelse distance objective > weapon-range [
    ifelse breed != targets [set last-heading heading
    face objective
    forward speed] [
    ]][
  if not [missile-inbound?] of objective [
    ifelse color = red [
      set red-missiles-counter red-missiles-counter + 1][
      set blue-missiles-counter blue-missiles-counter + 1]
    hatch-missiles 1 [
      set heading towards objective
      set speed missile-speed
      set size 4
      forward speed
      ]
      ask objective [
        set missile-inbound? true]]
  if is-HK? self or is-hunter? self [set weapons weapons - 1]
  set objective nobody
  if is-HK? self [set fly-back? true]
  set permission? false]
end

to get-permission
  ifelse [discernability] of objective + base-decision-time <= discernability-timer [
  set permission? true
  set discernability-timer 0][
  set discernability-timer discernability-timer + 1]
end

to resupply
  set resupplying? true
  ifelse xcor > min-pxcor + 1 and ycor > min-pycor + 1 and ((member? self killers and
(weapons = 0 or fuel <= 0)) or (member? self HKs and (weapons = 0 or fuel <= 0)) or
(member? self hunters and fuel <= 0))[
    facexy min-pxcor min-pycor
    forward min ( list speed distancexy min-pxcor min-pycor )] [
```

```
    ifelse resupply-timer < resupply-time [
      if resupply-timer = 0 and fuel = 0 [
        set refueling-counter refueling-counter + 1]
      if resupply-timer = 0 and weapons = 0 and (not member? self hunters) [
        set rearming-counter rearming-counter + 1]
      set color black
      set resupply-timer resupply-timer + 1] [
    ifelse return-last? and distancexy last-x last-y > 0.1 [
      set color blue - 2
      ifelse is-hk? self [set weapons max-hk-weapons
      set fuel max-hk-fuel] [
        ifelse is-killer? self [set weapons max-killer-weapons
          set fuel max-killer-fuel] [
          set fuel max-hunter-fuel]]
      facexy last-x last-y
      forward min (list speed (distancexy last-x last-y))][
    ifelse return-last? [
      set color blue - 2
      set resupply-timer 0
      set heading last-heading
      set resupplying? false] [
    set color blue - 2
    ifelse is-hk? self [set weapons max-hk-weapons
      set fuel max-hk-fuel] [
      ifelse is-killer? self [set weapons max-killer-weapons
        set fuel max-killer-fuel] [
        set fuel max-hunter-fuel]]
    set resupply-timer 0
    setxy min-pxcor + sensor-width / 2 min-pycor
    set heading 0
    set resupplying? false
    ]]]]
end

to call-in
  if any? (turtle-set (list killers hks)) with [weapons != 0 and objective = nobody] and not
[missile-inbound?] of objective and not any? ((turtle-set (list killers hks)) with [objective
= ([objective] of myself)])[
  compute-nearest-support
  if nearest-support-dist < comm-length [
    ask nearest-support [
    set last-heading-fb heading
    set last-x xcor
    set last-y ycor
    set objective [objective] of myself]
    set objective nobody
```

```
    set nearest-support nobody]]
end

to compute-nearest-support
  set nearest-support-dist distance min-one-of (turtle-set (list killers HKs)) with [weapons
!= 0 and objective = nobody] [distance myself]
  set nearest-support min-one-of (turtle-set (list killers HKs)) with [weapons != 0 and
objective = nobody] [distance myself]
end

to loiter
  ifelse random 100 < 50   [
    set heading heading + random 35] [
    set heading heading - random 35]
  forward speed
end

to fly-back
  set fly-back? true
  ifelse distancexy last-x last-y > 0.1 [
      facexy last-x last-y
      forward min (list speed (distancexy last-x last-y))][
      set heading last-heading-fb
      set fly-back? false]
end

;==============================================================
;==============================================================

to defenses
  ask targets with [armed != 0 and not hiding?] [
    ifelse objective != nobody [
    fire ][
    if (count HKs != 0 or count hunters != 0 or count killers != 0) and distance min-one-of
turtles with [breed != targets and breed != missiles] [distance myself] < sensor-range [
    set objective min-one-of turtles with [breed != targets and breed != missiles] [distance
myself]]
    ]]
end


;==============================================================
;Missile codes
;==============================================================

to missile-flight
```

```
  ask missiles [
    ifelse flight-time >= max-missile-flight [die] [
    ifelse [hiding?] of objective and random 100 > remember-percentage [
    ask objective [set missile-inbound? false]
    die] [
    ifelse distance objective < detonation-radius [
    kill] [
    ifelse abs (subtract-headings heading towards objective - 180) < 5 [
    kill] [
    set heading towards objective
    forward min ( list speed (distance objective) )
    set flight-time flight-time + 1
    ]]]]]
end

to kill
  ask objective [die]
  die
end

;===========================================================
;Target codes
;===========================================================
to target-move
  ask targets with [speed != 0 and not hiding?] [
    set heading heading + random 15
    forward speed]
end
to unhide
  ask targets with [hiding?] [
    if random 100 < popup-percent [
      set hiding? false
      set color red]]
  ask targets with [not hiding? and can-hide?] [
    if hiding-counter > loiter-time [
      set hiding? true
      set hiding-counter 0
      set color gray]]
  ask targets with [hiding? = false] [
    set hiding-counter hiding-counter + 1]
end

;===========================================================
;Data summarizing for output
;===========================================================
to finish
```

```
  file-open "test.txt"
    ;file-write "Threats ="
    file-write count targets with [variety = "SAM"]
    ;file-write ","
    ;file-write "Facilities ="
    file-write count targets with [variety = "facility"]
    ;file-write ","
    ;file-write "Trucks ="
    file-write count targets with [variety = "truck"]
    ;file-write ","
    ;file-write "HKs ="
    file-write count hks
    ;file-write ","
    ;file-write "Hunters ="
    file-write count hunters
    ;file-write ","
    ;file-write "Killers ="
    file-write count killers
    ;file-write ","
    ;file-write "Rearming ="
    file-write rearming-counter
    ;file-write ","
    ;file-write "Refueling ="
    file-write refueling-counter
    ;file-write ","
    ;file-write "Blue_Missiles ="
    file-write blue-missiles-counter
    ;file-write ","
    ;file-write "Red_Missile ="
    file-write red-missiles-counter
    ;file-write ","
    ;file-write "Ticks = "
    file-write ticks
  file-close
end
```

REFERENCES

[1] Artillery Battle. The NewsHour with Jim Lehr (Publich Broadcasting Service, 2002).

[2] RAH-66 Comanche Helicopter Production Continues. High Performance Composities (2003).

[3] U.S. Army Cuts Embattles Comanche Helicopter Program. High Performance Composities (2004).

[4] Webster's Unabridged Dictionary. In 2nd, ed (Random House, 2005).

[5] Monte Carlo Method. pp. This reference give a historical account of the development of Monte Carlo Analysis. It describes the principle of deriving an approximate solution to a quantitative problem using statistical means. It also provides some additional references, and mathematical basis for MCS. (Contingency Analysis, 2007).

[6] Anaylsis of Alternatives - Navy Discussion. In Defense Acquisition University, ed. Acquisition Community Connection (2007).

[7] David G. Gilmore, ed. Spacecraft Thermal Control Handbook (The Aerospace Press, El Segundo, CA, 2002).

[8] Abdi, H. A neural network primer. Journal of Biological Systems, 1994, 2(3), 247-283.

[9] aerospaceweb.org. Aicraft Museum - F-35 Lightening II. (2007).

[10] Ahn, B.-H., DeLaurentis, D. and Mavris, D.N. Advanced Personal Air Vehicle Concept Development Using Powered Rotor and Autogyro Configurations. AIAA Aircraft Technology, Integration, and Operations (ATIO) Forum (AIAA, Los Angeles, CA, 2002).

[11] Air Force Studies Board - National Research Council. Pre-Milestone A and Early-Phase Systems Engineering: A Retrospective Review and Benefits for Future Air Force Systems Acquisitions. (National Academies Press, Washington, D.C., 2008).

[12] Anderson, J. Fundamentals of Aerodynamics. (McGraw-Hill, New York 1991).

[13] Applied Physics Laboratory. Warfare Analysis Laboratory. (Johns Hopkins University, 2007).

[14] Army, U. Future Combat Systems (Brigade Combat Team) - White Paper. In Defense, D.o., ed (Department of Defense, 2007).

[15] Aseeri, A. Planning Under Uncertainty: Regret Theory. In Oklahoma, U.o., ed (2007).

[16] Associated Press. Engineer Says Saddam's Bunker Can Withstand Newest Bombs. USA Today (Berlin, 2003).

[17] Atkinson, R. Murky Ending Clouds Desert Storm Legacy. The Washington Post (Washington, DC, 1998).

[18] Baggesen, A. Design and Operational Aspects of Autonomous Unmanned Combat Aerial Vehicles. Modeling, Virtual Environments and Simulation (MOVES) (Naval Postgraduate School, Monterey, CA, 2005).

[19] Baker, A.P. and Mavris, D.N. Assessing the Simultaneous Impact of Requirements, Vehicle Characteristics, and Technologies During Aircraft Design. 39th AIAA Aerospace Sciences Meeting and Exhibit (AIAA, Reno, NV, 2001).

[20] BBC. 1988: US warship shoots down Iranian airliner. On This Day (BBC, 2007).

[21] Beal, C. Autonomous Weapons, Brave New World. Jane's Defence Weekly (2000).

[22] Bigelow, B. Global Hawk's soaring costs blasted: GAO assails Pentagon for high-risk strategy. San Diego Union-Tribune (San Diego, 2006).

[23] Biltgen, P.T., Ender, T., Mavris, D. N. Development of a Collaborative Capability-Based Tradeoff Environment for Complex System Architectures. 44th AIAA Aerospace Sciences Meeting and Exhibit (AIAA, Reno, Nevada, 2006).

[24] Biltgen, P.T. A Methodology for Capability-Based Technology Evaluation for Systems-of-Systems. Aerospace Engineering, p. 480 (Georgia Institute of Technology, Atlanta, 2007).

[25] Biltgen, P.T. and Mavris, D.N. A Technique for Interactive Probabilistic Multiple Attribute Decision Making. 45th AIAA Aerospace Sciences Meeting and Exhibit (AIAA, Reno, NV, 2007).

[26] Bin, A., Hill, R. and Jones, A. Desert Storm: A Forgotten War. (Greenwood Publishing Group, Westport, CT, 1998).

[27] Boeing Corporation. The B-1B Bomber. Backgrounder (The Boeing Corporation, St. Louis, MO).

[28] Borer, N. Decision Making Strategies for Probabilistic Aerospace Systems Design. Aerospace Engineering (Georgia Institute of Technology, Atlanta, GA, 2006).

[29] Bowman, J. Sensor Revisit Rate. (Air Force Research Lab, Dayton, OH, 2007).

[30] Bowman, J.C., Plumley, R.W., Dubois, J.A. and Wright, D.M. Mission Effectiveness Comparisons of Morphing and Non-Morphing Vehicles. 6th AIAA Aviation Technology, Integration and Operations Conference (ATIO) (AIAA, Wichita, Kansas, 2006).

[31] Bowman, J.C., Sanders, B., Cannon, B., Kudva, J., Joshi, S. and Weisshaar, T. Development of Next Generation Morphing Aircraft Structures. 48th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference (AIAA, Honolulu, HI, 2007).

[32] Bowman, S. Bosnia: U.S. Military Operations. (Congressional Research Service, Washington, DC, 2003).

[33] Box, G.E.P. and Draper, N.R. Empirical Model-Building and Response Surfaces. (John Wiley & Sons, New York, NY, 1987).

[34] Box, G.E.P. and Wilson, K.B. On the Experimental Attainment of Optimum Conditions (with discussion). Journal of the Royal Statistical Society, 1951, Series B 13(1), 1-45.

[35] Brand, S. James Dewar - "Long-term Policy Analysis". (The Long Now Foundation, 2004).

[36] Breyfogle, F.W.I. Implementing Six Sigma, Smarter Solutions Using Statistical Methods. (John Wiley & Sons, New York, NY, 1999).

[37] Bright, C.D. The Jet Makers: The Aerospace Industry from 1945 to 1972. (Regents Press of Kansas, 1978).

[38] British Broadcasting Corporation. Iraq Invades Kuwait. On This Day (BBC, London, UK, 2005).

[39] Brown, C.E. and Wiltse, J.C. Overview and Background. In Currie, N.C. and Brown, C.E., eds. Principles and Applications of Millimeter-Wave Radar (Artech House, Norwood, MA, 1987).

[40] Brown, D. Discussion on USAF Kill Chain. In Poole, B., ed (2008).

[41] Buxbaum, P.A. Extending the UAV. Military Aerospace Technology (2005).

[42] Calhoun, L. The Strange Case of Summary Execution by Predator Drone. Peace Review, 2003, 15(2), 209-214.

[43] Cameron, M.B. When Robots Kill. CSC 30 - Exercise New Horizons (Canadian Forces College, Toronto, Ontario, 2004).

[44] Canning, J.S. A Concept of Operations for Armed Autonomous Systems.

[45] Central Intelligence Agency. Iraq. The World Factbook (CIA, Langley, VA, 2008).

[46] Chief of Naval Operations. The United States Navy in "Desert Shield" / "Desert Storm". (Naval Historical Center, Washington, DC, 1991).

[47] Chona, R. Aerospace Structures for the 21st Century: An Air Force Perspective on Research Needs - Keynote Lecture. 46th AIAA/ASME/ASCE/AHS/ASC SDM ConferenceAustin, Texas, 2005).

[48] Clark, V. Sea Power 21: Projecting Decicisve Joint Capabilities. Proceedings (2002).

[49] Cochrane, C.B. Joint Program Management Handbook. (Defense Acquisition University Press, Fort Belvoir, VA, 2004).

[50] Committee for Mine Warfare Assessment. Naval Mine Warfare: Operational and Technical Challenges for Naval Forces. (National Academy Press, Washington, D.C., 2001).

[51] Conetta, C. and Knight, C. The Readiness Crisis of the U.S. Air Force: A Review and Diagnosis. Defense Alternatives Briefing Report #10 (The Commonwealth Institute, Cambridge, MA 1999).

[52] Croci, R. Radar Basics. (2002).

[53] Cyr, J. Space & Electronic Warfare Lexicon. Sarisota, FL, (2007).

[54] Danielle S. Soban and Mavris, D.N. The Need for a Military Systems Effectiveness Framework: The System of Systems Approach. 1st AIAA Aircraft, Technology, Integration, and Operations Forum (AIAA, Los Angeles, CA, 2001).

[55] Davidson, K. Sandstorms -- test of true grit. San Francisco ChronicleSan Francisco, 2003).

[56] Davis, J. The Cancellation of Crusader: A study in the Dynamics of Decision Making. USAWC Strategy Research Project (US Army War College, Carlisle Barracks, PA, 2003).

[57] Defense Acquisition University. Defense Acquisition Guidebook. (Defense Acquisition University Press, 2004).

[58] Defense Acquisition University. Integrated Defense Acquisition, Technology, & Logistics Life Cycle Management Framework (2005). (Defense Acquisition University Press, 2005).

[59] Defense Acquisition University. Command Briefing 2005/2006 In Defense, D.o., ed (Defense Acquisition University Press, 2005).

[60] Defense Industry Daily. Comanche's Child: The ARH-70 Armed Reconnaissance Helicopter. (2007).

[61] Defense Science Board. Defense Science Board Study on Unmanned Aerial Vehicles and Uninhabited Combat Aerial Vehicles. In Defense, D.o., ed (Washington, DC, 2004).

[62] Department of Defense. Systems Engineering Fundamentals. (Defense Acquisition University Press, Fort Belvoir, Virginia, 2001).

[63] Department of Defense. Joint Test and Evaluation Handbook. In Defense, D.o., ed (Defense Acquisition University, 2005).

[64] Deputy Undersecretary of the Navy (RD&A) Acquisition. Department of the Navy Acquisition and Capabilities Guidebook. (Defense Acquisition University, 2006).

[65] Dieter, G.E. Engineering Design. (McGraw-Hill Higher Education, USA, 2000).

[66] Dillon, W.R. and Goldstein, M. Multivariate Analysis. (John Wiley & Sons, Inc., New York, NY, 1984).

[67] Divine, M. NavySEALs.com (2004).

[68] Dombrowski, P. and Gholz, E. Buying Military Transformation: Technological Innovation and the Defense Industry. (Columbia University Press, New York, NY, 2006).

[69] Dudgeon, J.E. and Gopalakrishnan, R. Fractal-Based Modeling of 3D Terrain Surfaces. Southeastcon '96. 'Bringing Together Education, Science and Technology' (1996).

[70] Dunningan, J.F. and Nofi, A.A. Victory and Deceit: Dirty Tricks at War. (William Morrow and Company, New York, NY, 1995).

[71] Dupuy, T.N. Understanding Defeat: How to Recover from Loss in Battle to Gain Victory in War. (Nova Publications, McLean, VA, 1990).

[72] Durham, J. Capabilities Based Planning: An AT&L Perspective on FCB Interactions - PPT. (Office of the Under Secretary of Defense, Acquisition, Technology, and Logistics).

[73] Edwards, S.J.A. Swarming and the Future of Warfare. Public Policy Analysis (Pardee RAND Graduate School, Santa Monica, CA, 2004).

[74] Eggert, R.J. and Mayne, R.W. Probabilistic Optimal Design Using Successive Surrogate Probability Density Functions. Journal of Mechanical Design, 1993, 115, 385-391.

[75] Ehart, W. Simulated Annealing. (Sandia National Laboratory, 1997).

[76] Elliot, M.S. UAVs May Plan Increasing Operational Role. Air Force Print NewsWashington, 2003).

[77] Emch, G. and Parkinson, A.R. Using Engineering Models to Control Variability: Feasibility Robustness for Worst-Case Tolerances. ASME Design Automation Conference, p. 411 (ASME, Albuquerque, NM, 1993).

[78] Ender, T. A Top-Down, Hierarchical, System-of-Systems Approach to the Design of an Air Defense Weapon. Aerospace Engineering (Georgia Institute of Technology, Atlanta, 2006).

[79] Engelbrecht, L. Did software kill soldiers? IT Web (IT Web, Johannesburg, 2007).

[80] Engler, W.O., Biltgen, P.T. and Mavris, D.N. Concept Selection Using an Interactive Reconfigurable Matrix of Alternatives (IRMA). 45th AIAA Aerospace Sciences Meeting and Exhibit (AIAA, Reno, NV, 2007).

[81] Federation of American Scientists. Medium Altitude Endurance UAV, Predator. Advanced Concept Technology Demonstrations (ACTDs) Master Plan (Federation of American Scientists, 1996).

[82] Federation of American Scientists. F-111. (Federation of American Scientists, 1998).

[83] Federation of American Scientists. B-1B Lancer. WMD Around the World (Federation of American Scientists, 1999).

[84] Federation of American Scientists. M109A6 Paladin Self Propelled Howitzer. (Federation of American Scientists, 2000).

[85] Federation of American Scientists. M1 Abrams Main Battle Tank. Federation of American Scientists, 2000).

[86] Federation of American Scientists. AGM-114 Hellfire. (Federation of American Scientists, 2000).

[87] Feuchter, C.A. Air Force Analyst's Handbook: On Understanding the Nature of Analysis. (Office of Aerospace Studies, Kirtland AFB, NM, 2000).

[88] Fishman, C. Fresh Start 2002: Non-stop Flight. Fast Company.com, p. 82 (2001).

[89] Fogleman, R.R. Strategic Vision and Core Competencies. (Defense Issues: Volume 11, Number 96, 1996).

[90] Ford, R.B. and Barkan, P. Beyond Parameter Design - A Methodology Addressing Product Robustness at the Concept Formulation Stage. In Behun, J.R., ed. National Design Engineering Show and Conference, pp. 1-7 (ASME, Chicago, Illinois, 1995).

[91] Foster-Miller Inc. Products & Service: TALON Military Robots, EOD, SWORDS, and Hazmat Robots. (2007).

[92] Franklin, D. and Burroughs, J. NCDC: Climate of Iraq. In Commerce, D.o., ed (National Oceanographic and Atmospheric Administration, 2007).

[93] Frederick, M. UAVs Take to the Skies. Space News Business Report (2006).

[94] Friedman, S.M. The Inflation Calculator. (2007).

[95] Fulghum, D.A. and Wall, R. Hypersonic Weapons Attack Time Problem. Aviation Week and Space Technology, p. 30 (2000).

[96] Garcia-Ortiz, A. Application of the Analytical Hierarchy Process to the Design of a User-Adaptable, Tactical Decision Aid. Canadian Conference on Electrical and Computer Engineering, pp. 595 - 598 (IEE, 1994).

[97] General Accounting Office. Operation Desert Storm: Evaluation of the Air Campaign. (Report to the Ranking Minority Member, Committee on Commerce, House of Representatives, Washington, DC, 1997).

[98] General Accounting Office. Air Force Operating and Support Cost Reductions Need Higher Priority. (2000).

[99] Globalsecurity.org. B-52 G. In Pike, J., ed. Weapons of Mass Destruction (WMD) GlobalSecurity.org (GlobalSecurity.org, 2005).

[100] Globalsecurity.org. C-130 Hercules. In Pike, J., ed. GlobalSecurity.org (GlobalSecutity.org, 2005).

[101] Globalsecurity.org. Operation Desert Shield/Desert Storm Timeline. In Pike, J., ed. Military (GlobalSecurity.org, 2005).

[102] Globalsecurity.org. Operation Desert Storm. In Pike, J., ed. Military (GlobalSecurity.org, 2005).

[103] Globalsecurity.org. B-1. In Pike, J., ed. Weapons of Mass Destruction (WMD) GlobalSecurity.org (GlobalSecurity.org, 2005).

[104] Globalsecurity.org. F-111 Aardvark. In Pike, J., ed. Military GlobalSecurity.org (GlobalSecurity.org, 2006).

[105] Globalsecurity.org. RAH-66 Comanche. (Global Security.org 2007).

[106] Globalsecurity.org. M109A6 Paladin Self Propelled Howitzer. In Pike, J., ed. Military (GlobalSecurity.org, 2008).

[107] Globalsecurity.org. BGM-109 Tomahawk. In Pike, J., ed. Military (GlobalSecurity.org, 2008).

[108] Goebel, G. Unmanned Aerial Vehicles. (2007).

[109] Graybill, F. Theory and Application of the Linear Model. (Duxbury Press, North Scituate, MA, 1976).

[110] Guetlein, M.A. Lethal Autonomous Weapons - Ethical and Doctrinal Implications. Joint Military Operations (Naval War College, Newport, RI, 2005).

[111] Guevara, E. Guerrilla Warfare: Introduction by Marc Becker. (University of Nebraska Press, Lincoln, NE, 1961).

[112] Hammons, S. Navy SEALs and Marines use positive human traits and virtues for success. American Chronicle (Ultio LLC, 2007).

[113] Hanlon, E.J. and Route, R.A. Enhanced Networked Seabasing. (Navy Warfare Development Command, Newport, RI, 2002).

[114] Hernandez, G., Simpson, T.W., Allen, J.K., Bascaran, E., Avila, L.F. and Salinas, F. "Robust Design of Product Families with Production Modeling and Evaluation." Journal of Mechanical Design. 2001, 123(2), 183-190.

[115] Hernandez, L.M., Durch, J.S., Blazer, D.G. and Hoverman, I.V. Gulf War Veterans: Measuring Health. (National Academies Press, 1999).

[116] Hobbins, T. Using Information Technology to Win the Nation's War. Intercom, pp. 4-5 (2004).

[117] Hoffren, J. and Sailaranta, T. Maneuver Autopilot for Realistic Performance Model Simulations. AIAA Modeling and Simulation Technologies Conference and Exhibit (AIAA, Montreal, Canada, 2001).

[118] Howard, W., Pilling Donald. Defense Science Board Task Force on Sea Basing. (Office of the Under Secretary of Defense for Acquisition, Technology and Logistics, 2003).

[119] Hughes-Wilson, J. Military Intelligence Blunders and Coverups. (Carroll & Graf Publishers, New York, NY, 2004).

[120] Hutchison, H.C. Super Hornet Shoots Down F-35s. Dirty Little Secrets (StrategyWorld.com, 2007).

[121] Huyse, L. Solving Problems of Optimization Under Uncertainty as Statistical Decision Problems. 42nd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference and Exhibit (AIAA, Seattle, WA, 2001).

[122] Inspector General. Report on the Audit of the Acqisition of the Tacit Rainbow Anti-Radiation Missile System. In Defense, D.o., ed (Department of Defense, Arlington, VA, 1991).

[123] International Council on Systems Engineering. INCOSE Systems Engineering Handbook, v3.1. (INCOSE Central Office, Seattle, WA, 2007).

[124] Israeli Aerospace Industries. Harpy. (2002).

[125] Jaiswal, N.K. Military Operations Research: Quantitative Decision Making. (Kluwer Academic Publishers, Boston, MA, 1997).

[126] Jimeno, J. and Mavris, D.N. An Automated Robust Process for Physics Based Aerodynamic Prediction. 2000 World Aviation Conference (AIAA, San Diego, CA, 2000).

[127] Johnson, C. and Schutte, J. BRAINN 2.0 Users Manual. (ASDL, Atlanta, GA, 2006).

[128] Joint Chiefs of Staff. Joint Operations. (DoD, Washington, DC, 2008).

[129] Joint Doctrine Division J-7 Joint Staff. DoD Dictionary of Military Terms. (Department of Defense, Washington, DC, 2007).

[130] Kayne, R.A. AoA Case Study: A Variation of Regret Analysis. 73rd MORS Symposium (MORS, Alexandria, VA, 2005).

[131] Kazmer, D. and Roser, C. Evaluation of Product and Process Design Robustness. Research in Engineering Design, 1999, 11(1), 20-30.

[132] Keane, A.J. Statistical Improvement Criteria for Use in Multiobjective Design Optimization. AIAA Journal, 2006, 44(4), 879-891.

[133] Keaney, T. and Cohen, E. Gulf War Air Power Survey. (US Government Printing Office, Washington, DC, 1993).

[134] Kennedy, M. KC-135 Recapitalization Analysis of Alternatives Overview - ARSAG. 2005).

[135] Kennedy, M. and al, e. Analysis of Alternatives (AoA) for KC-135 Recapitalization. (RAND Corporation, Arlington, VA, 2006).

[136] Kirby, M.R. A Methodology for Technology Identification, Evaluation, and Selection in Conceptual and Preliminary Aircraft Design. Aerospace Engineering (Georgia Insitute of Technology, Atlanta, GA, 2001).

[137] Kirby, M.R., Raczynski, C., Mavris, D. An Approach for Strategic Planning of Future Technology Portfolios. (AIAA, 2006).

[138] Kopp, C. Smart Tankers: Hubs in the Networked Force. DefenceToday Magazine2005).

[139] Kotz, N. Wild Blue Yonder: Money, Polotics, and the B-1 Bomber. (Princeton Press, Princeton, New Jersey, 1988).

[140] Kovach, C. and Gauthier, J.S. A Missile System Simulation Benchmark Using the ACSL Graphic Modeller. AIAA Modeling and Simulation Technologies Conference and Exhibit (AIAA, Montreal, Canada, 2001).

[141] Lempert, R.J., Popper, S. W., Bankes, S. C. Shaping the Next One Hundred Years: New Methods for Quantitative, Long-Term Policy Analysis. (The RAND Corporation, Santa Monica, CA, 2003).

[142] Lempert, R. Scenario Analysis Under Deep Uncertainty. DOE/EPA Workshop on Modeling the Oil Transition (2006).

[143] Lennox, D. Jane's Air-Launched Weapons. (Jane's Information Group, Surrey, 2007).

[144] Lewe, J.-H., Byong-Ho-Ahn, Delaurentis, D.A., Mavris, D.N. and Schrage, D.P. An Integrated Decision-Making Method to Identify Design Requirements through Agent-Based Simulation for Personal Air Vehicle System. AIAA's Aircraft Rechnology, Integration, and Operations (ATIO) 2002 Technical (AIAA, Los Angeles, CA, 2002).

[145] Lewis, L. and Parkinson, A. Robust Optimal Design using a Second-Order Tolerance Model. Research in Engineering Design, 1994, 6, 25-37.

[146] Loeb, V. Rumsfield Untracks Crusader. Washington Post (Washington, D.C., 2002).

[147] Lopez, T.C. F-22 Excels at Establishing Air Dominance. Air Force Print News (2006).

[148] Lowry, R.S. The Gulf War Chronicles: A Military History of the First War with Iraq. (iUniverse, Lincoln, NE, 2003).

[149] Lussier, M. An Analysis of U.S. Army Helicopter Programs. In Office, T.C.B., ed (The United States Congress, 1995).

[150] Maier, M.W. Architecting Principles for Systems of Systems. Sixth Annual Insternational Symposium, International Council on Systems Engineering (INCOSE, Boston, MA, 1996).

[151] Mankins, M.C. Technology Readiness Levels. (NASA, 1995).

[152] Massey, K., Heiges, M., DiFrancesco, B., Ender, T. and Mavris, D. A System-of-Systems Design of a Guided Projectile Mortar Defense System. 24th Applied Aerodynamics Conference (AIAA, San Francisco, CA, 2006).

[153] Mattingly, J., Heiser, W. and Pratt, D. Aircraft Engine Design. (AIAA, Reston, VA, 2002).

[154] Mavris, D. AE 6373 Class Notes. (Georgia Institute of Technology, Atlanta, 2005).

[155] Mavris, D.N., Bandte, O., Schrage, D.P. Robust Design Simulation: A Probabilistic Approach to Multidisciplinary Design. AIAA Journal of Aircraft 1999, 36(1), 298-307.

[156] Mavris, D., Biltgen, P. and Weston, N. Advanced Design of Complex Systems Using the Collaborative Visualization Environment (CoVE). 43rd AIAA Aerospace Sciences Meeting and Exhibit (AIAA, Reno, NV, 2005).

[157] Mavris, D.N., DeLaurentis, D.A., Bandte, O., Hale, M.A. A Stochastic Approach to Multi-disciplinary Aircraft Analysis and Design. (AIAA, 1998).

[158] Mavris, D.N., DeLaurentis, D.A. A Stochastic Design Approach for Aircraft Affordability. 21st Congress of the International Council on the Aeronautical Sciences (ICAS) (ICAS, Melbourne, Australia, 1998).

[159] Mavris, D.N., Bandte, O. and Schrage, D.P. Application of Probabilistic Methods for the Determination of an Economically Robust HSCT Configuration. 6th AIAA, NASA, and ISSMO Symposium on Multidisciplinary Analysis and Optimization (AIAAA, Reston, VA, 1996).

[160] Mavris, D.N. and Kirby, M.R. Preliminary Assessment of the Economic Viability of a Family of Very Large Transport Configurations. (AIAA, 1996).

[161] McCarthy, J.D. Seabasing Logistics. NDIA 10th Annual Expeditionary Annual Expeditionary Warfare Conference (2005).

[162] MIT Lincoln Laboratory. DARPA Artificial neural network Study. (MIT Lincoln Laboratory, Lexington, MA, 1998).

[163] Moniz, D. Pentagon Reports Fewer US Casualties. USA Today (2005).

[164] Moore, F.G. Approximate Methods for Weapon Aerodynamics. (AIAA, Reston, VA, 2000).

[165] Morris, G.V. Airborne Pulsed Doppler Radar. (Artech House, Norwood, MA, 1988).

[166] Moshkina, L. and Arkin, R.C. Lethality and Autonomous Systems: Survey Design and Results. (Georgia Institute of Technology, Atlanta, GA, 2007).

[167] Mount, M. Troops Put Thorny Questions to Rumsfeld: Defense Chief Speaks to Iraq-bound Soldiers in Kuwait (CNN, Atlanta, GA, 2004).

[168] Myers, R. and Montgomery, D. Response Surface Methodology. (John Wiley & Sons, New York, 2002).

[169] National Museum of the USAF. Boeing B-52G Stratofortress. Fact Sheets (USAF, 2006).

[170] National Museum of the USAF. General Dynamics F-111F Aardvark. Fact Sheets (USAF, 2007).

[171] National Museum of the USAF. Northrop AGM-136A TACIT RAINBOW. Fact Sheets (USAF, 2007).

[172] Office of Aerospace Studies. AoA Handbook. In Defense, D.o., ed (2004).

[173] Office of Aerospace Studies. AoA Online Course. (USAF, 2006).

[174] O'Halloran, J. Jane's Land-Based Air Defence 2006-2007. (Jane's Information Group, Surrey, UK, 2006).

[175] Oklahoma City Air Logistics Center. AGM-86B Security Classification Guide. (Federation of American Scientists, 1997).

[176] O'Rourke, R. Navy-Marine Corps Amphibious and Maritime Prepositioning Ship Programs: Background and Oversight Issues for Congress. (Congressional Research Sevices, Washington, DC, 2006).

[177] Orr, M. Introduction to Radial Basis Function Networks. (University of Edinburgh, Edinburgh, Scotland, 1996).

[178] Park, G.-J., Lee, T.-H., Lee, H.H. and Hwang, K.-H. Robust Design: An Overview. AIAA Journal, 2006, 44(1), 181-191.

[179] Patel, C., Kirby, M.R. and Mavris, D. Niched-Pareto Genetic Algorithm for Aircraft Technology Selection Process. 11th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference (AIAA, Portsmouth, VA, 2006).

[180] Phoenix Integration. Model Center 7.0. (Blacksburg, VA, 2006).

[181] Pianta, D. Measures of Effectiveness (MoE) and Measures of Performance (MoP). (University of Queensland, Queensland, Au, 2001).

[182] Pike, J. RQ-4A Global Hawk (Tier II+ HAE UAV). Intelligence Resource Program (Federation of American Scientists, 1998).

[183] Pinker, A., Smith, C.G. and Booher, J.W. Selecting Effective Acquisition Process Metrics. Acquisition Review Quarterly, 1997, 189-208.

[184] Poole, J. Tactics of the Crescent Moon: Militant Muslim Combat Method. (Posterity Press, Emerald Isle, NC, 2004).

[185] Purcell, V. The Boxer Uprising: A Background Study. (Cambridge University Press, Cambridge, 1963).

[186] Pyzdek, T. The Six Sigma Handbook: The Complete Guide for Greenbelts, Blackbelts, and Managers at All Levels. (McGraw-Hill, New York, NY, 2003).

[187] RAND Corporation. Alternatives for Recapitalizing the U.S. Air Force KC-135 Aerial Refueling Tanker Fleet - Summary Briefing. 2006).

[188] Rapp, S. Creating a Modeling and Simulation Baseline. In Cloud, D.J. and Rainey, L.B., eds. Applied Modeling and Simulation: An Integrated Approach to Development and Operation (McGraw-Hill, New York, NY, 1998).

[189] Raymer, D.P. Aircraft Design: A Conceptual Approach. (AIAA, Reston, VA, 1999).

[190] Raytheon Corporation. Tomahawk Cruise Missile. (Ratheon, Tuscon, Arizona, 2006).

[191] Renze, J. and Weisstein, E. Analysis - from Mathworld. (Wolfram, 2004).

[192] Ricks, T.E. Rumsfeld Gets Earful From Troops. Washington Post (Washington, DC, 2004).

[193] Ritchey, T. General Morphological Analysis: A general method for non-quantified modelling. 16th EURO Conference on Operational Analysis (Sweedish Morphological Society, Brussels, 2006).

[194] Roskam, J. Airplane Design Part VII: Airplane Cost Estimation: Design Development, Manufacturing, and Operation. (DARcorporation, Lawrence, KS, 2002).

[195] Rowland, S.C. Enhanced Networked Sea Basing. Expeditionary Warfare Conference (DTIC, 2002).

[196] SAS Institute Inc. JMP Statistics and Graphics Guide. (SAS, Cary, NC, 2005).

[197] SAS Institute Inc. Fly High in Your Opportunity SPace. JMP Customer Brief (SAS Institute,, Cary, NC, 2006).

[198] Sato, J., Shah, N. and Bergman, S. Strike Unmanned Aerial Vehicle (UAV) Innovative Concept Exploration. In DoD, ed (Air Force Research Lab, 2005).

[199] Sato, J., Shah, N. and Bergman, S. Mission Area Assessment in Support of Future Long Term Challenges. In DoD, ed (Air Force Research Lab, 2005).

[200] Scharl, J. and Mavris, D. Building Parametric and Probabilistic Dynamic Vehicle Models Using Neural Networks. AIAA Modeling and Simulation Rechnologies Conference and Exhibit (AIAA, Montreal, Canada, 2001).

[201] Schetzen, M. Airborne Doppler Radar: Applications, Theory, and Philosophy. (AIAA, Reston, VA, 2006).

[202] Schrage, D.P. Army Aircraft Requirements in the 1990's: A View Forward. AIAA Aircraft Design, Systems, and Technology Meeting (AIAA, Fort Worth, TX, 1983).

[203] Schwartz, N.A. "Joint Capabilities Integration and Development System". In Defense, D.o., ed (2005).

[204] Secretary of The Air Force. USAF Intelligence Targeting Guide. In Defense, D.o., ed (1998).

[205] Shachtman, N. Inside the Robo-Cannon Rampage (Updated). (wired.com, 2007).

[206] Shao, T., Krishnamurty, S. and Wilmes, G.C. Preference-Based Surrogate Modeling in Engineering Design. AIAA Journal, 2007, 45(11), 2688-2701.

[207] Shin, S., Guo, Y., Choi, Y. and Choi, M. Development of a Robust Data Mining Method Using CBFS and RSM. (Department of Systems Management Engineering, Inje University, Gimhae, Kyung-Nam 621-749, South Korea, 2006).

[208] Siegner, J.J. Analysis of Alternatives: Multivariate Considerations. (Air Force Institute of Technology, Dayton, OH, 1998).

[209] Signor, Davis and al, e. Efficient Air Traffic Scenario Generation.

[210] Singh, J. Of War and Destruction. (The Hindu Folio 2000).

[211] Singley, G.T., Smith, R.L. and Schrage, D.P. Army Light Family of Rotorcraft (LHX) Concept Formulation. AIAA Aircraft Design, Systems, and Technology Meeting (AIAA, Fort Worth, TX, 1983).

[212] Smith, D.J. Reliability, Maintainability, and Risk. (Butterworth-Heinemann, Great Britain, 2001).

[213] Soban, D. A Methodology for the Probabilistic Assessment of System Effectiveness as Applied to Aircraft Survivability and Susceptibility. Aerospace Engineering (Georgia Institute of Technology, Atlanta, GA, 2001).

[214] Sproles, N. Establishing Measures of Effectiveness for Comand and Control: A Systems Engineering Perspective. In Defence, D.o., ed (Defence Science & Technology Organisation, 2001).

[215] Stanfill, K. IPPD. (University of Florida, Gainsville, FL, 2008).

[216] Streetly, M. Jane's Radar and Electronic Warfare Systems 2006-2007. (Jane's Information Group, Surrey, UK, 2006).

[217] Stulberg, A. “Managing the Unmanned Revolution in the U.S. Air Force,” Orbis 51:2 (Spring 2007), pp. 251-265.

[218] TACOM. Recapitalization of the US Army. Warren, Michigan.

[219] Taguchi, G. Taguchi on Robust Technology Development. (ASME Press, New York, 1993).

[220] Taleb, N.N. The Black Swan: The Impact of the Highly Improbable. (Random House, New York, NY, 2007).

[221] Taylor, J. Jane's All the World Aircraft, 1989-90. (Jane's Information Group, Surrey, UK, 1990).

[222] Ternion Corporation. "FLAMES Example Model Manual, Version 6.1". (Ternion Corporation, 2007).

[223] Ternion Corporation. "FLAMES Application Manual, Version 6.1". (Ternion Corporation, 2007).

[224] Ternion Corporation. "FLAMES Sensor Coverage Option Manual, Version 6.0". (Ternion Corporation, 2005).

[225] Thomas, V. A Method for Scenario-based Risk Assessment for Robust Aerospace Systems. Aerospace Engineering (Georgia Institute of Technology, Atlanta, GA, 2007).

[226] Tirpak, J.A. Find, Fix, Track, Target, Engage, Assess. Journal of the Air Force Association, 2000, 83(7).

[227] Twiss, B.C. Forecasting for Technologists and Engineers: A practival guide for better decisions. (Peter Peregrinus Ltd., London, 1992).

[228] Tyson, A.S. 2006 Budget: Agency Breakdown. The Washington PostWashington, D.C., 2005).

[229] Tzu, S. The Art of War. (Oxford University Press, London, UK, 1963).

[230] Under Secretary of Defense for Acquisition Technology and Logistics. Mandatory Procedures for Major Defense Acquisition Programs (MDAPS) and Major Automated Information System (MAIS) Acquisition Programs. In Defense, D.o., ed (National Technical Information Service, 2001).

[231] United States Marine Corps. Officers Candidate School - Candidate Regulations. In Defense, D.o., ed (2003).

[232] United States Marine Corps. Officers Candidate School - Course Notes. In Defense, D.o., ed (2003).

[233] US Army. Paladin. US Army Fact File (US Army,, 2007).

[234] US Army. Army Technology - Crusader 155mm - Self Propelled Howitzer. (SPG Media Limited 2007).

[235] US Navy. Tomahawk Cruise Missile. United States Navy Fact File (2007).

[236] USAF. The USAF Intelligence Targeting Guide. In Defense, D.o., ed (1998).

[237] USAF. The U.S. Air Force Remotely Piloted Aircraft and Unmanned Aerial Vehicle Strategic Vision. In Defense, D.o., ed (2005).

[238] USAF. Global Hawk. US Air Force Factsheets (USAF, 2006).

[239] USAF. C-130 Hercules. US Air Force Factsheets (USAF, 2006).

[240] USAF. B-52 Stratofortress. US Air Force Factsheets (USAF, 2007).

[241] USAF. B-1B Lancer. US Air Force Factsheets (USAF, 2007).

[242] USAF. KC-135 Stratotanker. US Air Force Factsheets (USAF, 2007).

[243] Vanderplaats, G.n. Numerical Optimization Techniques for Engineering Design. (Vanderplaats Research and Development, Inc., Colorado Springs, CO, 2001).

[244] Vedantam, S. Tolerance for a War's Death Toll Depends on How You Look at It. Washington Post (Washington, DC, 2006).

[245] Vest, J. Fourth-generation Warfare. The Atlantic (2001).

[246] Vogt, A. and Bared, J.G. Accident Models for Two-Lane Rural Roads: Segments and Intersections. In Office of Safety and Traffic Operations R&D, F.H.A., ed (Federal Highways Administration, 1998).

[247] Walsh, D.C. Will unmanned aerial vehicles run out of air? Government Computer News (2006).

[248] Wang, W., Beeson, D., Wiggs, G. and Rayasam, M. A Comparison of Meta-modeling Methods Using Practical Industry Requirements. 47th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference (AIAA, Newport, RI, 2006).

[249] Wei, Z., Bao-chun, L. and Hui-zhong, W. Modeling and Simulation Approach for Multi-disciplinary Virtual Prototyping. 4th World Congress on Intelligent Control and Automation (IEEE, Shanghai, P.R. China, 2002).

[250] Weisstein, E.W. ANOVA. Mathworld- A Wolfram Web Resource (2007).

[251] Weisstein, E.W. "Minimax Approximation." MathWorld--A Wolfram Web Resource (2008).

[252] Wheelan, J. Jefferson's War: America's First War on Terror, 1801–1805. (Carroll & Graf, New York, 2003).

[253] White, J. LHX Platform Technology Assessment. In Army, D.o.t., ed (1983).

[254] Wilcox, J. Weapons Technology Blueprint for the Future. Precision Strike Winter Roundtable (DoD, 2007).

[255] Wilde, D.J. Product Quality in Optimization Models. In Taylor, D.L., ed. 4th International Conference on Design Theory and Methodology, pp. 237-241 (ASME, Scottsdale, Arizona, 1992).

[256] Williams, C. Introduction. In Williams, C., ed. Holding the Line: U.S. Defese Alternatives for the Early 21st Century (The MIT Press, Cambridge, MA, 2001).

[257] Wills, G. Technolological Forecasting: The Art and its Managerial Implications. (Penguin, Baltimore, 1972).

[258] Wolf, F. Introduction to the Scientific Method. (Rochester, NY, 2006).

[259] Womack, J.P., Jones, D.T. and Roos, D. The Machine that Changed the World. (Free Press, 2007).

[260] Xiong, Y., Weichen, Apley, D. and Ding, X. A Nonstationary Covariance Based Kriging Method for Metamodeling in Engineering Design. 11th AIAA/ISSMO Multidiciplinary Analysis and Optimization Conference (AIAA, Portsmouth, VA, 2006).

[261] Yost, D.C. The Future of U.S. Overseas Presence. Joint Forces Quarterly, pp. 70-82 (1995).

[262] Zink, P.S., Raveh, D.E. and Mavris, D.N. Robust Structural Design of an Active Aeroelastic Wing with Maneuver Load Inaccuracies. Journal of Aircraft, 2004, 41(3), 585 - 593.

[263] Zipfel, P.H. and O'Grady, T. Missile Model Fidelity and Integration for Air Combat Simulators. (AIAA, 1999).

VITA

Benjamin Hancock Poole

Benjamin Poole was born in Columbus, Georgia. He attended Lassiter High School in Marietta, GA, received a B.S. in Aerospace Engineering with Honor from the Georgia Institute of Technology, Atlanta, Georgia in 2004 and an M.S. in Aerospace Engineering from the Georgia Institute of Technology in 2005. As an undergraduate, Mr. Poole served in the USMC PLC program and interned as a structural engineer at Naval Air Weapons Station China Lake. While in graduate school, he worked as a graduate researcher in the Aerospace Systems Design Laboratory in the School of Aerospace Engineering and in the Aerospace, Transportation, and Advanced Systems Laboratory at the Georgia Tech Research Institute on numerous projects for military and intelligence community customers. He also served as a teaching assistant in the School of Aerospace Engineering and a guest lecturer for systems engineering in the Distance Learning and Professional Education program. In 2005, Mr. Poole led Georgia Tech's first place effort for the AIAA Graduate Team Aircraft Design Competition. He also participated in the 2007-2008 Sam Nunn Security Fellowship Program, where he studied the international and national security implications of autonomous, persistent-strike unmanned aerial vehicles.