2021-06

METHODOLOGY FOR MODELING COST AND SCHEDULE RISK ASSOCIATED WITH RESOURCE DECISIONS INVOLVING THE U.S. ARMY'S MODERNIZATION EFFORTS FOR 2035

McClary, Matthew A.

Monterey, CA; Naval Postgraduate School

http://hdl.handle.net/10945/67775

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

METHODOLOGY FOR MODELING COST AND SCHEDULE RISK ASSOCIATED WITH RESOURCE DECISIONS INVOLVING THE U.S. ARMY'S MODERNIZATION EFFORTS FOR 2035

by

Matthew A. McClary

June 2021

Thesis Advisor:	Gregory K. Mislick
Co-Advisors:	Alejandro S. Hernandez
	Brian M. Wade
Second Readers:	Matthew G. Boensel
	Jenifer R. McClary

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC, 20503.			
1. AGENCY USE ONLY ***(Leave blank)***	**2. REPORT DATE** June 2021	**3. REPORT TYPE AND DATES COVERED** Master's thesis	
4. TITLE AND SUBTITLE METHODOLOGY FOR MODELING COST AND SCHEDULE RISK ASSOCIATED WITH RESOURCE DECISIONS INVOLVING THE U.S. ARMY'S MODERNIZATION EFFORTS FOR 2035			**5. FUNDING NUMBERS**
6. AUTHOR(S) Matthew A. McClary			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A			**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release. Distribution is unlimited.			**12b. DISTRIBUTION CODE** A

13. ABSTRACT (maximum 200 words)

Prioritization decisions using the Army Modernization and Analysis (AMA)-developed Trade-Space Decision Exploration System (TRADES) does not address programmatic variance related to cost and schedule growth. This study offers an improved methodology for modeling cost risk by employing sound cost estimation principles, distribution fitting, Monte Carlo simulations, and cost/benefit analysis to assist strategic decision makers and the acquisitions community. To that end, this approach follows a five-step methodology that (1) collects and screens cost data from the Cost Assessment Database Enterprise (CADE), (2) determines normalized cost growth factors, (3) identifies and constructs the appropriate distributions for modeling, (4) simulates cost variance among the entire program portfolio, and (5) recommends the necessary contingency cash reserve quantity associated with a decision maker's confidence level. The result is a credible, repeatable, and effectual cost estimating methodology that promotes commodity-based models for predicting cost growth and measuring risk.

14. SUBJECT TERMS cost estimation, cost variance, risk management, Monte Carlo simulation, lognormal distribution, Army Modernization and Analysis, AMA, Trade-Space Decision Exploration System, TRADES, Cost Assessment Database Enterprise, CADE			**15. NUMBER OF PAGES** 103
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

METHODOLOGY FOR MODELING COST AND SCHEDULE RISK ASSOCIATED WITH RESOURCE DECISIONS INVOLVING THE U.S. ARMY'S MODERNIZATION EFFORTS FOR 2035

Matthew A. McClary
Major, United States Army
BS, U.S. Military Academy, 2010

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN SYSTEMS ENGINEERING

from the

NAVAL POSTGRADUATE SCHOOL
June 2021

Approved by: Gregory K. Mislick
Advisor

Alejandro S. Hernandez
Co-Advisor

Brian M. Wade
Co-Advisor

Matthew G. Boensel
Second Reader

Jenifer R. McClary
Second Reader

Ronald E. Giachetti
Chair, Department of Systems Engineering

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Prioritization decisions using the Army Modernization and Analysis (AMA)-developed Trade-Space Decision Exploration System (TRADES) does not address programmatic variance related to cost and schedule growth. This study offers an improved methodology for modeling cost risk by employing sound cost estimation principles, distribution fitting, Monte Carlo simulations, and cost/benefit analysis to assist strategic decision makers and the acquisitions community. To that end, this approach follows a five-step methodology that (1) collects and screens cost data from the Cost Assessment Database Enterprise (CADE), (2) determines normalized cost growth factors, (3) identifies and constructs the appropriate distributions for modeling, (4) simulates cost variance among the entire program portfolio, and (5) recommends the necessary contingency cash reserve quantity associated with a decision maker's confidence level. The result is a credible, repeatable, and effectual cost estimating methodology that promotes commodity-based models for predicting cost growth and measuring risk.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

ACAT	Acquisition Category
A-D	Anderson-Darling
AFC	Army Futures Command
AMA	Army Modernization Analysis
AMSAA	Army Material Systems Analysis Activity
ANOVA	analysis of variance
ASL	Army senior leaders
BY	base year
CADE	Cost Assessment Data Enterprise
CAIV	cost as an independent variable
CAPE	Cost Assessment and Program Evaluation
CCDR	contractor cost data report
CE	cost estimation
CFT	cross-functional team
CGF	cost growth factor
Chi^2	Chi-Squared
CSDR	cost and software data report
CSRUH	Cost Schedule and Risk Uncertainty Handbook
CV	coefficient of variation
DOD	Department of Defense
E-IBCT	Early Infantry Brigade Combat Team
ICEAA	International Cost Estimating and Analysis Association
INCOSE	International Council of Systems Engineering
JSF	Joint Strike Fighter
K-S	Kolmogorov-Smirnov (K-S)
KT	key technology
LRIP	low-rate initial production
MAIMS	money assigned is money spent
MDAP	major defense acquisition program
MS	milestone

NPS	Naval Postgraduate School
OE	operational effectiveness
OSD	Office of the Secretary of Defense
PCTL	percentile
PDF	probability density function
PE	percent error
RDT&E	research, development, testing and evaluation
ROV	Real Options Valuation
SAR	Selected Acquisition Report
SD	standard deviation
SPAR24	Strategic Portfolio Analysis Review 2024
TPLN	three-parameter log-normal
TRAC	The Research and Analysis Center
TRADES	Trade-Space Decision Exploration System
TRL	technology readiness level
TY	then year
VIBES	Visual Interactive Beta Estimation System
WBS	work breakdown structure
WSMR	White Sands Missile Range

EXECUTIVE SUMMARY

As history has proven, cost overrun in defense spending is nearly inescapable; therefore, risk management is a necessity. Currently, the AMA estimation approach captures cost data based on a deterministic approach that lacks consideration for inevitable cost and schedule overruns that plague more than 80% of defense industries (Smart 2021, xi). This thesis, however, proposes a methodology that captures financial risk associated with historical defense programs based on commodity or milestone. This research leverages a statistical approach to capture the probability of cost overrun based on historical CADE data to inform decision makers on program rankings and contingency cash reserve levels necessary to achieve a desired confidence level.

The result is a quantifiable recommended cash reserve that supports the decision maker's desired confidence level for maintaining cost and schedule objectives without sacrificing technical performance or operational effectiveness. Findings include program comparisons to identify the elements of the portfolios that contribute the most risk based on coefficients of variation. By mitigating subjectivity through data-driven distributions and improving foresight via Monte Carlo simulations, this thesis bridges the gap between specific program uncertainty and industry trends to develop an objective CE methodology that adequately informs investment decisions.

This thesis employs an analytical approach to investigate how historical cost data can inform cost growth predictions based on commodity and nearest milestone. The process begins with gathering relevant historical CADE data before calculating each program's cost growth factor (CGF). The CGF data leads to the construction of probability distributions used to model cost and schedule behaviors. Next, this study captures the uncertainty by leveraging Monte Carlo simulations to generate probability plots for total program costs. In doing so, this method can determine the appropriate cash reserve quantity necessary to meet a decision maker's confidence level for staying underbudget. Analysis of variance tests provide the mechanism for testing the trends of variance between factors to identify the most effective method for predicting cost growth. Figure 1 demonstrates the overall process for determining these forecasted portfolio costs.

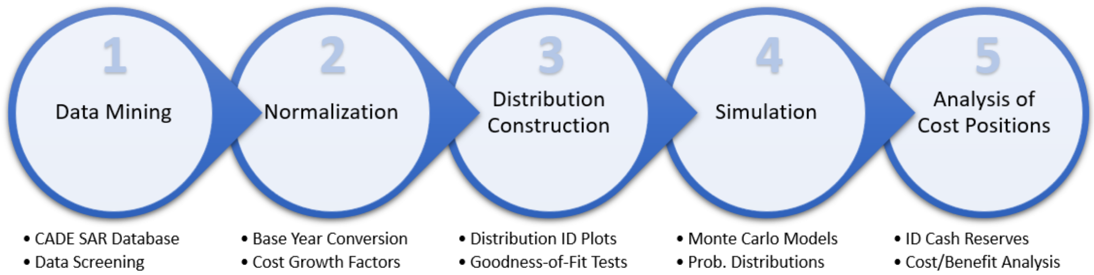


Figure 1. Thesis Methodology Process Flowchart

Based on the statistical analysis performed in this study and throughout the proposed methodology, evidence supports the following insights and recommendations:

- There is no statistical difference in mean CGF values between commodities, milestones, or their subsets. Therefore, *a single benchmark CGF provides relatively rapid and effective cost growth insight when under time constraints*.

- The variance within programs provides insight into their inherent risk while *coefficients of variation provide the metric for prioritizing risk levels between programs or portfolios*.

- Cost variance peaks in milestone B and specific commodities (rotary), therefore, *analysts can counter the assumption that increased cost overrun before low-rate initial production (LRIP) does not necessarily imply that the program's cost risk is escalating out of control. Further investigation into specific CSDR submissions can provide insight into the reasons for cost overrun.*

- Forecasting cost using the milestone approach requires subjective assumptions for determining event events of developmental programs (AMA). Consequently, this study recommends *the commodity-focused cost estimating methodology since it promotes conservatism and objectivity*.

- Analysis of historical CADE data reveals that the lognormal distribution is the best model for cost growth. When coupled with Monte Carlo simulation techniques, *cost-prediction simulations provide a sound mechanism for translating confidence levels into contingency cash reserve quantities*.

- When "exceptional variation" (Smart 2021) is present in the simulated results for total portfolio cost, the recommended funding is at or above the 80% confidence level, but *an appropriate confidence level for funding can rest somewhere between 50% and 90% depending on risk tolerance and resource availability. This mitigates the risk of costly budget interruptions that hinder schedule and technical performance.*

Implementation of these recommendations would allow cost analysts to provide a structured approach to informing budget decisions and program prioritization. The simplicity of a single benchmark CGF value manages the expectations of senior leaders while analysts provide the due diligence that accounts for historical trends. Leveraging CADE data also reduces subjectivity while streamlining the estimating process since SME-elicitation is unnecessary. Diversification of portfolios helps to mitigate the increased risk in particular commodities but relies on the ability to shift resources between programs. Lastly, modeling and simulation provides the means for quantifying the risk and unpredictability that is intrinsic to government contracting. Overall, cost growth is undeniable, so, cost estimators have an obligation to capture and communicate that truth to the relevant decision makers. The methodology outlined in this thesis drives current practice closer toward the goal of accurate and precise cost prediction.

Reference

Smart, Christian B. 2021. *Solving for Project Risk Management: Understanding the Critical Role of Uncertainty in Project Management.* New York: McGraw-Hill Education.

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

This is dedicated to my magnificent wife, Jenifer, for her unwavering support and extraordinary skills, and to my incredible daughters, Madison and Finley, for their inspiring energy and endless enthusiasm.

THIS PAGE INTENTIONALLY LEFT BLANK

PERMISSION

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION TO RESEARCH

A. PURPOSE

This thesis aims to develop a methodology that captures financial risk associated with defense programs based on commodity and milestone. Military leaders must exercise prudence when determining how to steward government resources toward achieving operational effectiveness. Fundamentally, senior leaders weigh their decisions based on risk versus reward. This thesis focuses on the risk aspect of decision-making with emphasis on monetary investments in Army modernization programs. The Department of the Army relies on the newly established Army Futures Command (AFC) to prioritize and oversee the development of all modernization programs. Currently, AFC models consider risk and benefit based on a deterministic approach that lacks consideration for inevitable cost and schedule overruns that plague more than 80% of defense industries (Smart 2021, xi). This research leverages a statistical approach to capture the probability of cost overrun to inform decision makers on appropriate program rankings and contingency cash reserve levels necessary to achieve the desired operational effectiveness and meet budgetary constraints.

B. PROBLEM BACKGROUND

In the wake of the 2018 *National Defense Strategy*, military and civilian leadership established AFC to address modernization gaps due to an erosion of close combat capabilities relative to threats around the world. As a result, AFC began investing in eight cross-functional team (CFT) modernization programs to achieve strategic military advantage by 2035. This rapid change in organizational structure and investment strategy requires new a methodology for adequately understanding, analyzing, and informing program investments of the future.

In preparation for potential budget constraints within the Department of Defense, AFC must accurately prioritize more than 80 modernization programs based on cost versus operational effectiveness. To accomplish the prioritization feat, AFC relies on a series of strategic research and analysis centers located throughout the continental United States; one of which resides at the Naval Postgraduate School (NPS). The Research and Analysis

Center (TRAC) in Monterey leverages the intellectual capital of faculty and students to achieve their mission of conducting "relevant and credible applied research toward improving military operations analysis" (Wade 2020, under "Our Mission").

In support of AFC's effort to establish program prioritization within the Army's portfolio, TRAC-Monterey, in conjunction with their higher headquarters in Kansas, provides analytic evidence for establishing an integrated approach known as Army Modernization Analysis (AMA) (Luher et al. 2021). The AMA team recently designed an analytic tool for the AFC CG to rapidly assess the cross-portfolio effects of trades between programs with respect to the cost and operational benefit. TRAC refers to the resulting methodology to as the "Trade-Space Decision Exploration System" (TRADES) tool. The intent of TRADES is to inform investment decisions made by Army senior leaders (ASL) impacted by resource and time constraints by analyzing the changes to cost and operational benefit (Luher et al. 2021).

Currently, prioritization decisions using the AMA-developed TRADES method does not address programmatic uncertainty associated with cost and schedule; it also provides a false sense of certainty based on deterministic modeling and point estimates. The ever-present uncertainty of defense contracting presents significant risk once military decision makers choose to invest in that specific program over another. In this case, the current prioritization model does not capture the risk associated with the probability of cost overrun. Furthermore, the TRADES tool does not include a reliable methodology for estimating the cost to cancel a program that considers industry, schedule, procurement, or Research, Development, Test, and Evaluation (RDT&E). This thesis provides a methodology for modeling cost and schedule-related risk to inform AFC's prioritization efforts.

The authoritative source of cost data is the Cost Assessment Data Enterprise (CADE) managed by the Office of the Secretary of Defense Cost Assessment and Program Evaluation (OSD CAPE) department. CADE's mission is "to increase analyst productivity and effectiveness by collecting, organizing, and displaying data in an integrated single web-based application, improving data quality, and reporting compliance and source data transparency" (CADE 2020, under "About CADE"). The CADE website asserts itself to being the single authority in providing joint source data that is accurate and easily researchable. It recently replaced and upgraded the original CAPE database named Defense Cost and Resource Center (DCARC).

C. PROBLEM STATEMENT

The Army Modernization Analysis (AMA) Team of The Research and Analysis Center (TRAC) requires an estimation methodology for total program cost that considers schedule risk to inform modernization decisions relevant to the Army Futures Command (AFC).

D. THESIS OBJECTIVES AND DELIVERABLE

This thesis employs an analytical approach to investigate how historical cost data can improve cost estimation methodology. The study incorporates quantitative analyses, which include computational statistical methods of analysis. The overall research will address the following objectives:

1. Construct distributions for total program cost based on historical industry and technology maturation data.
2. Leverage Monte Carlo simulations to capture variance and confidence levels associated with cost by industry.
3. Offer data-driven information during program selection.
4. Support development of cost positions in the context of operational effectiveness.

The ultimate deliverable of this research and analysis is an improved methodology for modeling cost risk. This study includes a proof-of-concept that compares the model's results with historical programmatic cost data as a means of cross-validation. The intent is to employ the lessons learned from this study when providing future recommendations to the commander of AFC.

E. THESIS SCOPE, LIMITATIONS, AND ASSUMPTIONS

To develop a holistic understanding of data trends, this study scopes the problem by focusing on Army programmatic commodities relative to their nearest acquisition milestone. The three commodities included in this study are (1) vehicles, (2) rotary systems, and (3) C3I (command, control, communication, and intelligence) programs. Scoping the research to three specific platforms allows this study to concentrate on insights related to developing a repeatable methodology for capturing cost and schedule risk.

This thesis will limit the cost data to that which comes from the CADE website. The historical data for each corresponding program comes from the Cost Assessment Data Enterprise system maintained by the Office of the Secretary of Defense Cost Assessment and Program Evaluation (OSD CAPE) directorate. Within the CADE database resides two repositories. The first database includes detailed cost data on every line-item within a single program while the second captures macro-level summary data for an entire program. This study will focus on the latter for consistency since TRAC-provided cost data only includes top-level metrics.

Additionally, the CADE database only captures cost data for programs that satisfy the criteria for major defense acquisition programs (MDAP) and Acquisition Category Level-I (ACAT-I). Government officials classify programs as MDAP or ACAT-I based on a cost threshold or congressional oversight requirements. Fortunately, cost variation based on program size is relatively consistent across the spectrum but reporting requirements due to congressional oversight typically correlate with programs that suffer from severe cost overrun (Smart 2020). This dynamic thereby introduces the potential to overestimate cost growth thereby leading to a more conservative estimate when applying the approach proposed by this research.

According to TRAC-AMA, the Army G-8 and AFC directorate are responsible for drafting cost and schedule data connected to each of the modernization programs (Luher et al. 2021). However, these cost estimates only consider equipping costs and do not include sustainment and personnel costs, since the AFC commander’s authority only extends to equipping. As such, this study chose historic programs that shared similar cost estimates.

The following assumptions apply to this thesis:

1. AMA-provided data reflects point estimates from Contractor Cost Data Reports (CCDR).
2. All AMA modernization programs are developmental technology despite whether they are improvements to legacy systems.
3. AMA-provided cost data reflects then-years dollars based on a universal inflation rate.
4. Programs that transitioned/restructured before reaching low-rate initial production (LRIP) are inadmissible due to their falsely perceived cost-savings.
5. Selected acquisition reports (SAR) that fall within six months of a milestone event are redundant and therefore excused from analysis.
6. The earliest SAR acts as the milestone A (MS-A) report since many programs are not required to submit a SAR until MS-B.

F. OVERVIEW OF LITERATURE REVIEW

To comprehend the existence and magnitude of cost and schedule risk, one must understand the factors that lead to overspending. Figure 2 illustrates the many internal and external contributors that lead to cost and schedule overrun based on the literature review of project risk management experts. Highlighted at the end of each vector, there are common problematic categories listed in bold with the corresponding programmatic risk factors annotated in the parentheses. The tags along each diagonal of the fishbone diagram in Figure 2 illustrate additional contributors within that category. For example, Christian Smart, author of *Solving for Project Risk Management*, emphasizes the Department of Defense's frivolous tendency to consider funding as "money assigned is money spent" (MAIMS), which becomes a mechanism (highlighted in orange) for overspending. Additionally, he asserts that the defense acquisition environment (mother nature) typically follows Parkinson's Law which states that "work expands to fill the time available," and Hofstadter's Law: "It always takes longer than you expect, even when you take into account Hofstadter's Law" (Smart 2021, 9). As a result, over 80% of defense programs suffer from cost and schedule overrun (2021). This thesis will primarily focus on the "method" and "measurement" aspects of the fishbone diagram in Figure 2.

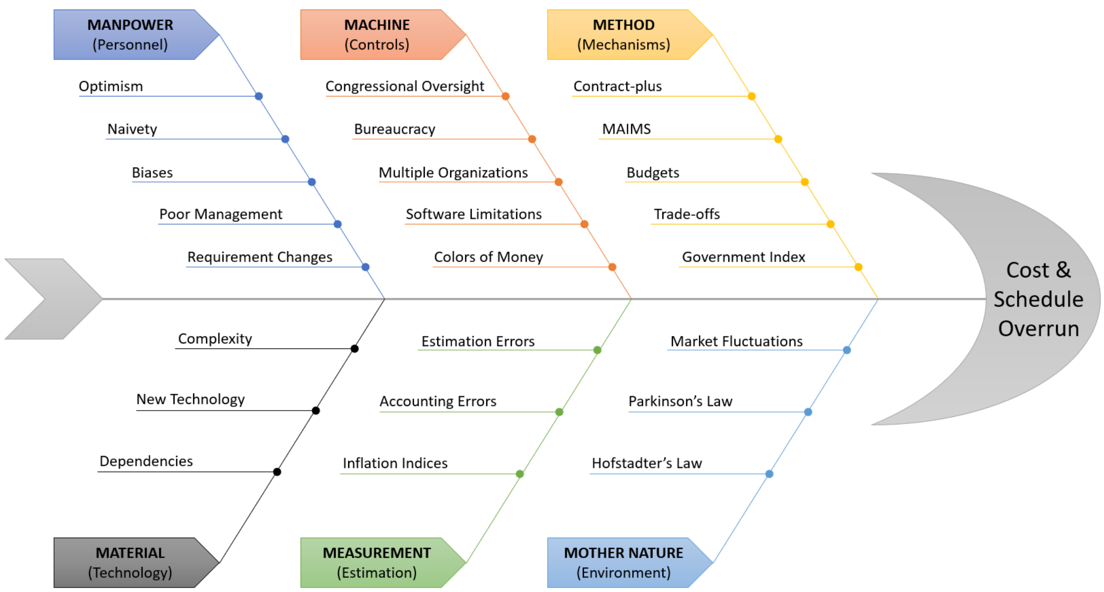


Figure 2. Fishbone Diagram: Causes of Cost and Schedule Overrun

G. DEFINITIONS

The *Joint Agency Cost Schedule and Risk Uncertainty Handbook* (CSRUH) defines risk, opportunity, and uncertainty in the following terms:

- Risk is the probability of a loss or injury.
- Opportunity is a favorable event or outcome.
- Uncertainty is the indefiniteness about the outcome of a situation. (Department of Defense [DOD] and National Aeronautics and Space Administration [NASA] 2014)

Figure 3 illustrates the context of each variable with respect to cost estimation.

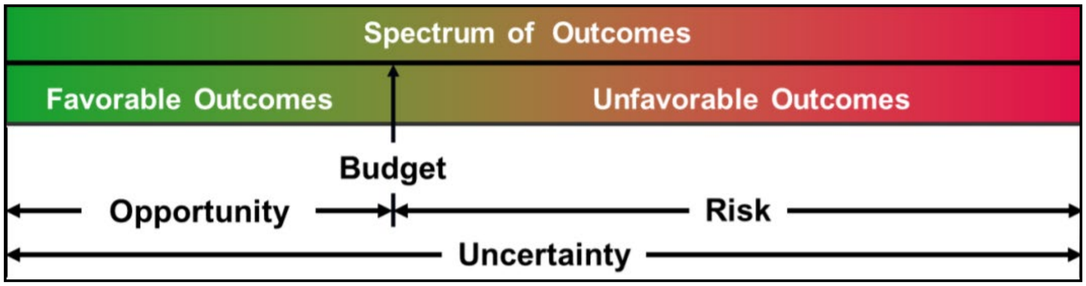


Figure 3. Risk, Opportunity, and Uncertainty in the Context of Cost Estimation. Source: DOD and NASA (2014).

H. OVERVIEW OF METHODOLOGY

The methodology for accomplishing the research objectives relies on gathering historical data from the military's Cost Assessment Data Enterprise (CADE) database; this data leads to the construction of probability distributions that will serve as models for cost and schedule behaviors based on the specific commodity and nearest milestone. Next, this study captures the uncertainty associated with total program cost by leveraging Monte Carlo simulations. The subsequent output provides the data for determining cash reserve levels necessary to meet a decision maker's confidence levels. Analysis of variance tests provide the mechanism for testing the trends of variance between factors to identify the most effective method for predicting cost growth. The probability of cost overrun serves as a metric for capturing potential risk; these inferences then inform program prioritization when compared with TRAC's assessed operational effectiveness.

This study leverages multiple simulation software platforms found within Excel to verify cost modeling efforts. Using a subsection of the historical data from CADE, the k-fold cross-validation technique acts as the primary mechanism for validating the improved methodology.

I. BENEFITS OF THESIS

The benefit of this work provides strategic decision makers and the acquisitions community a means to incorporate historically based confidence levels in cost estimation. Analyzing the historical data of past and present Department of Defense programs can provide invaluable insight into the cost-spirals that torment budgeting and prioritizing efforts across all branches of service. Too often, decision makers rely on point estimates developed by biased organizations that fail to capture the cost and schedule risk inherent to defense contracting.

This study provides a repeatable methodology using any pool of historical data to conceptualize cost trends across industries, focus risk mitigation efforts, and prepare for future investments that maximize operational effectiveness. The resulting benefit is a comprehensive, credible, and well-documented method for articulating the risk associated with acquisition-related decisions (Mislick and Nussbaum 2015).

The findings of this thesis can lend more credibility to the systems engineering (SE) enterprise by employing sound statistical analysis and modeling to mitigate risk and promote the qualitative aspect of a systems engineering approach. Capturing cost and schedule-related risk is one of the greatest challenges in SE; it is often ill-defined within reports, yet it represents a major component of decision analysis. In the realm of resource constraints, cost and schedule risk is frequently underappreciated and ignored when compared to technical risk (Smart 2020). By replicating the methodology presented in this work, systems engineers can improve their understanding of risk mitigation efforts related to cost estimates.

J. ORGANIZATION OF THESIS

This thesis follows the Naval Postgraduate School's Systems Engineering department writing guide for research work. As such, and to provide context to the problem, Chapter II provides a literature review of relevant and credible sources on the topics of cost estimation, program management, cost and schedule risk, Monte Carlo simulation, and methodology development for risk-informed analysis in acquisition. Chapter III highlights the methodology, process, and conditions for determining cost and schedule risk using simulations configured with distributions developed from CADE's historical data. Chapter IV employs the methodology, validates the model, and analyzes the data against AMA's estimates to develop findings. Lastly, Chapter V presents insights and meaning to the analytical work while highlighting unaddressed areas and topics for future research.

II. THEORETICAL FRAMEWORK AND LITERATURE REVIEW

This literature review covers four thematic topics that discuss relevant and credible research on modeling program risk. The four themes include (1) understanding cost and schedule risk, (2) risk analysis methodologies, (3) risk modeling and simulation, and (4) Army Modernization Analysis (AMA). The first three themes cover existing studies while the fourth encompasses how this thesis introduces new insights on programmatic and portfolio risk based on the parametric relationships between commodity and acquisition milestone events. Each section identifies how existing literature contributes to the thesis objectives outlined in Chapter I.

A. UNDERSTANDING COST AND SCHEDULE RISK

This section of the literature review starts by emphasizing the inevitable truth of cost and schedule overrun that plagues the government contracting industry. Areas of focus include (1) the need for risk management and (2) the ability to integrate cost and schedule risk.

1. The Need for Risk Management

In support of his recent book on program risk management, Christian Smart (2021) analyzed the total cost of 289 Department of Defense and NASA programs before concluding that the average cost growth equaled 52%. In addition, the proportion of programs that experienced cost overruns exceeded 80% with more than 90% of them experiencing schedule delays. Even from the opposite perspective, former Lockheed Martin CEO, Norman Augustine, corroborated Smart's (2021) findings in what he called the "Las Vegas Factor of Development Program Planning" where he claimed that the average increase for cost of development programs was also 52%.

As a result of cost overrun, program managers often sacrifice technical performance to maintain budget integrity which inherently degrades the combat effectiveness of military organizations once those systems arrive at the user level. One of the most effective ways of mitigating cost overrun is to simply focus on efforts for holding contractors accountable through good business practices. Yet, all too often, managers tend to focus their attention on cost savings opportunities rather than managing and mitigating the overwhelming cost consequences that plague more than 80% of defense programs (Smart 2021). By recognizing, measuring, and planning for cost-related risks, leaders can better ensure program success and thereby improve military modernization efforts.

Although Smart presents some compelling statistics on overall cost growth, his analysis concludes at the macro-level of insight by only delineating between major industries such as NASA, DOD, and civil engineering programs. This thesis, however, identifies program specific cost risk based on parametric attributes involving commodity and milestone events. The net benefit is a more useful tool for understanding, budgeting, and selecting programs based on risk tolerances of the decision maker. This effort coincides with the third research objective of this thesis: offer data-driven information during program selection.

2. Integrating Schedule Risk

During the 2002 INCOSE International Symposium, authors David Hulett and Bill Campbell composed an article on how to model cost and schedule risk in a single Monte Carlo Simulation. Their model presents a compelling method for introducing schedule risk when modeling cost while providing insight into how industries create risk based on the triple constraints of project management: cost, schedule, and performance. Figure 4 provides an illustration of the combined effects of cost and schedule growth versus the initial estimate at completion (EAC). As seen, the mean of the cumulative distribution for total project cost is substantially higher when compounding both cost and schedule risk, thus revealing the danger of underestimation should cost be the only factor of project risk analysis (Hulett and Campbell 2002).

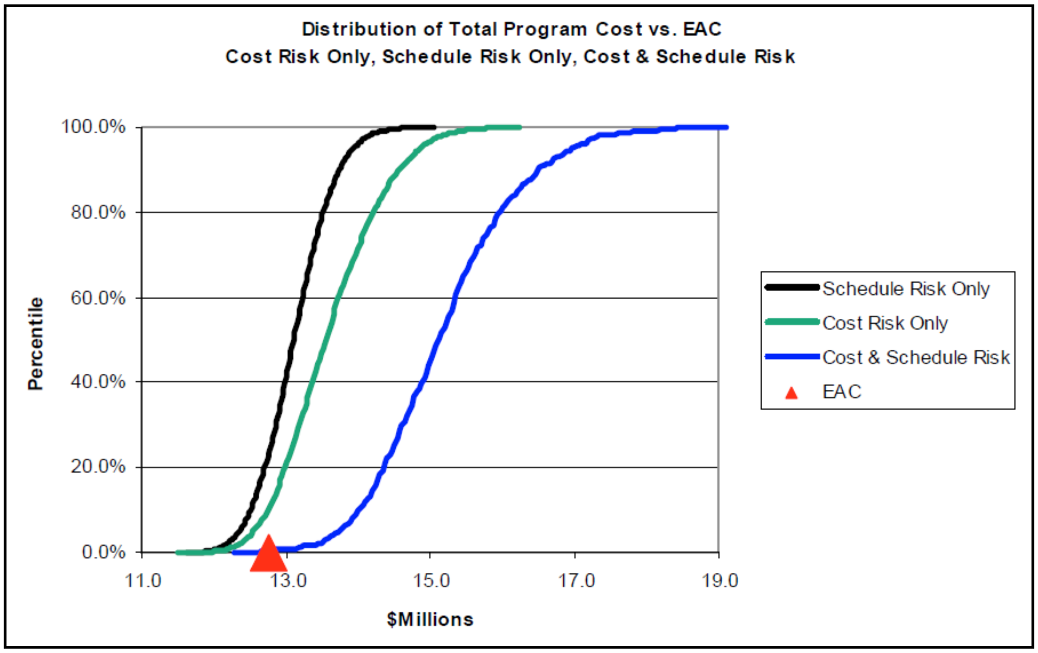


Figure 4. Cost and Schedule Risk Elements Combined.
Source: Hulett and Campbell (2002).

Hulett and Campbell's work highlights the importance of separating independent costs and variable costs when modeling cost risk and addresses the concept of a "contingency reserve" used by project cost estimators. Their research emphasizes the need to analyze project variance to inform estimators rather than relying on the historic percentage-based engineering practices. The result is a single cost-based model that incorporates schedule delays using cost-per-unit of time to determine the probability distribution of total project cost. This thesis leverages a similar approach for determining the probability of cost overrun based on cost-integrated schedule delays. Additionally, the concept of "contingency cash reserves" echoes Smart's recommendation for providing quantifiable risk management measures. Chapters IV and V of this study apply the insights of Smart, Hulett, and Campbell to address the programmatic and portfolio risk associated with current Army modernization programs.

B. RISK ANALYSIS METHODOLOGIES

This section of the literature review addresses the past and present attempts to measure, quantify, and respond to the seemingly inexorable reality of cost overrun. Areas of focus include (1) the qualitative and quantitative approach to risk management, (2) the pitfall of integrating expert advice, and (3) RAND's systems approach to capturing risk from 2015.

1. Qualitative and Quantitative Approach

Author K. Kansala (1997) is a principal consultant with the Nokia Research Center and subject matter expert on software cost estimation. In his article titled "Integrating Risk Assessment with Cost Estimation," Kansala presents a tool that draws from "questionnaires and project history to help calculate project risk contingencies" in the software industry (1997, 61). Kansala's method provides an exhaustive approach to dissecting quantifiable risk activities based on probability, magnitude, and outcome based on qualitative and quantitative responses from industry experts. However, the author admits that his method is subjective due to its dependency on input probabilities. Figure 5 illustrates the method by which Kansala elicits qualitative (subjective) responses in the bolded "definition of probability" section before determining the quantitative magnitude of impact values in the shaded "analysis" section along the top of the graphic.

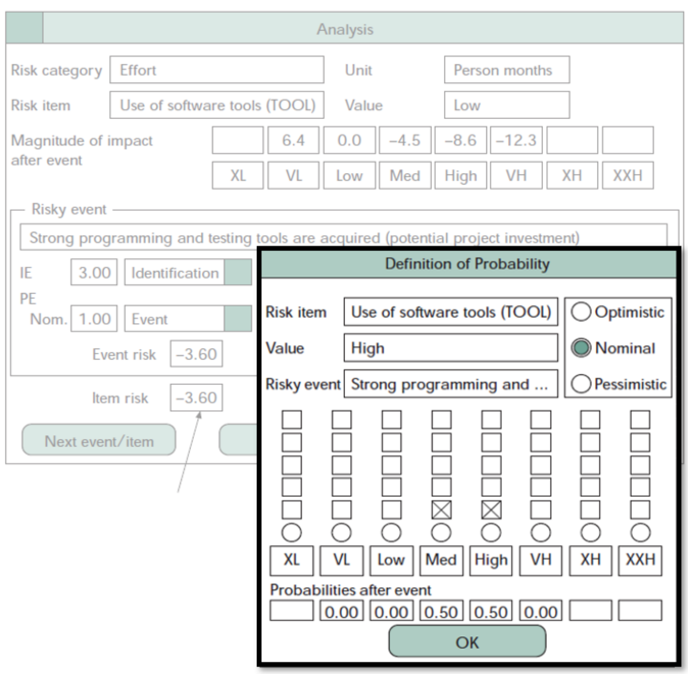


Figure 5. Risk Analysis Input Screen. Source: Kansala (1997).

Ultimately, Kansala's research illuminates the need to derive generally accepted industry standards on probability distributions for cost variance based on historical data rather than subject matter expert (SME) advice which can often include heuristic bias. This motive supports the first objective of this thesis: construct distributions for total program cost based on historical data. In doing so, it is possible to reduce potential biases and provide objective recommendations to the decision-maker at hand.

2. Integrating Expert Advice

Frederick W. Raymond (1999) is a retired government service employee with 30 years of experience in spacecraft acquisition and procurement management. His article titled "Quantify Risk to Manage Cost and Schedule" provides a process-oriented method to risk management that leverages quantifiable expert judgment for developing relevant triangular distributions that model risk uncertainty (1999). Raymond's research highlights the dichotomy of expert advice being the "crux" of analysis but also the weakest focus on the risk management process. Although dated, Raymond demonstrates a traditional method for creating a triangular distribution using the minimum, most likely, and maximum expected costs based on expert advice. Figure 6 illustrates the basic structure of the triangular distribution method.

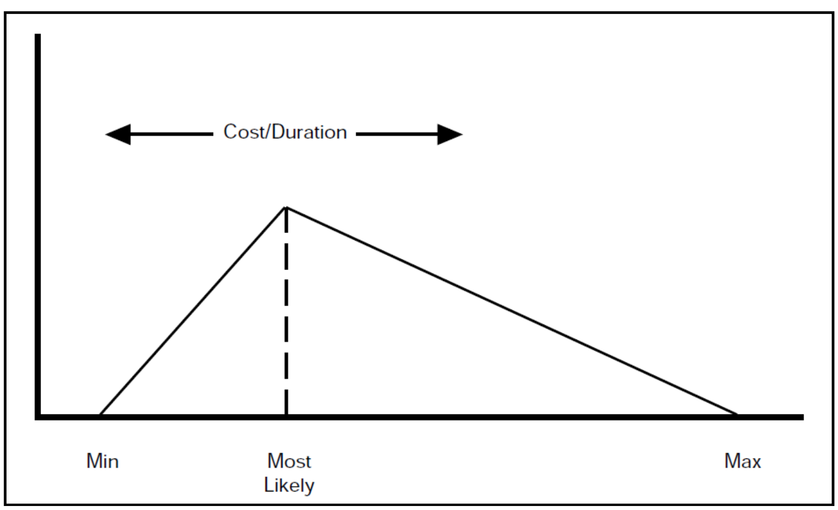


Figure 6. Risk Uncertainty for Cost or Duration. Source: Raymond (1999).

Raymond attempts to mitigate potential bias by statistically quantifying risk factor multipliers associated with categorical levels of overall program risk (low, moderate, high, and very high). He then uses the multipliers to modify the expert's estimated best-case scenario to determine the apexes of the triangular distribution. Table 1 depicts the proposed cost risk factor multipliers developed by Raymond based on his extensive experience in project management, evaluation, and estimation. The minimum cost factor multiplier for all programs regardless of risk level is the initial point estimate, therefore, the table indicates a multiplier of one. The subsequent factors (most likely and maximum) increase in value to represent the increased cost overrun as the risk level rises.

Table 1. Risk Factor Multipliers for Developing Triangular Distribution. Source: Raymond (1999).

Code		**Min**	**Most Likely**	**Max**
Low	L	1	1.04	1.10
Low+	L+	1	1.06	1.15
Moderate	M	1	1.09	1.24
Moderate+	M+	1	1.14	1.36
High	H	1	1.20	1.55
High+	H+	1	1.30	1.85
Very high	V	1	1.46	2.30
Very high+	V+	1	1.68	3.00

Of note, the factors express exponential growth as risk or variance increase in magnitude. Figure 7 illustrates the multiplier growth over the spectrum of risk from low to very high (not to exact scale). The lightly shaded area represents the "most likely" factor while the dark area illustrates the maximum cost multiplier for each category of risk. Raymond's concept of risk growth relates with this study's focus on the relationship between technology maturity and cost variance.

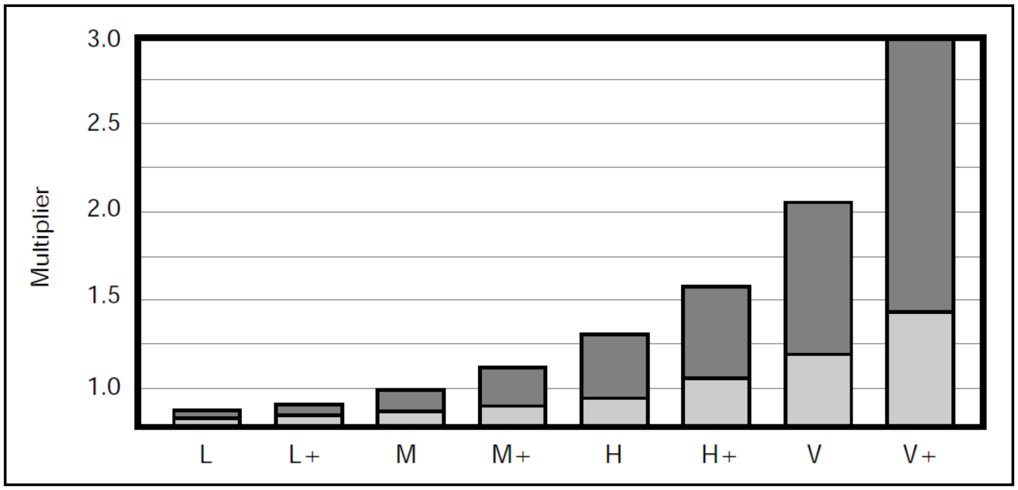


Figure 7. Graphical Representation of Risk Factors. Source: Raymond (1999).

However, the author admits that the risk factor multipliers merely provide an interim example that requires further investigation into historical data before publishing an industry-wide standard. Raymond's research emphasizes the need to analyze large-scale data that was once inaccessible. Fortunately, the newly established CADE database that this thesis leverages provides a unique opportunity to exploit industry-wide cost data toward developing a generally accepted solution for new risk factor multipliers based on historical data.

3. RAND Corporation Methodology Example

In a more recent study, the RAND Corporation conducted a comprehensive research project titled, *Developing a Methodology for Risk-Informed Trade-Space Analysis in Acquisition* to create a risk assessment tool that captures the quantifiable uncertainty associated with cost, schedule, and performance trade-offs (Bond et al. 2015). RAND achieved its objective by dissecting each program into key technological (KT) components before assigning stochastic properties to each based on SME-elicited times, costs, and performance values. In conjunction with the U.S. Army Material Systems Analysis Activity (AMSAA) Risk Team, the RAND Corporation published a 162-page report on how and why the new methodology could illuminate the consequences of cost, schedule, and performance trade-offs. The following 10 items summarize RAND's sequential steps for developing a program's risk profile:

1. Identify critical technology within each alternative.
2. Determine a schedule for each KT component.
3. Determine the consequence for failed KT.
4. Choose a critical path that mitigates KT consequences.
5. Define the cost-schedule relationship.
6. Draw a stochastic schedule date for each KT component.
7. Calculate a schedule estimate by aggregating schedule dates.
8. Calculate a performance estimate by aggregating consequences.
9. Calculate a cost estimate by aggregating KT costs associated with delivery dates.
10. Obtain distributions for cost, schedule, and performance.

Interestingly, much like the examples before, the RAND methodology also relies on expert opinion rather than quantifiable evidence. Their report highlights the fact that all their risk workshop data was based on subjective judgments from SMEs that "cannot be validated, nor have experts been calibrated to ensure some stability or realism in their opinions" (2015, 82). Moreover, the RAND Excel-based model proved to be hardcoded for triangular distributions powered only by the SME-elicited data. As a recommendation for further research, RAND suggests the development of distributions built from historical data (thesis objective #1). Their report leaves the reader with a desire to investigate the benefit of leveraging new and quantifiably based distributions for modeling cost and schedule growth.

C. RISK MODELING AND SIMULATION

This section of the literature review concentrates the research into the critical factors of modeling cost variance (risk). Key points of emphasis speak to (1) the need to identify realistic probability distributions, (2) the importance of accurate cost estimation techniques, and (3) the value of Monte Carlo simulations.

1. Probability Distributions

Although the triangular distribution is most common when modeling risk, there exists counter arguments in its ability to capture realistic activity costs or durations. By nature of its design, the triangular distribution terminates at its farthest endpoints and therefore fails to capture the extreme outliers and variance that sporadically occur. As such, it is incapable of forecasting rare occurrences like the Joint Strike Fighter (JSF) program where actual cost doubled the original estimate. Unless the program is definitively bounded against a cost ceiling, using the triangular distribution (Figure 6) will often underestimate risk (Smart 2021).

Furthermore, the linear shape of the triangular spread supports an overemphasis on the tails of the distribution while neglecting to capture the full probability for values around the most likely scenario (Kuhl et al. 2009). In an article written by several esteemed professors across multiple engineering disciplines titled, "Introduction to Modeling and Generating Probabilistic Input Processes for Simulation," the authors (Kuhl et al. 2009) suggest the use of the Beta distribution in lieu of the triangular. They argue that the natural curve of the Beta distribution is more like the Gaussian distribution and therefore better represents the realistic phenomenon of random activities. As seen in Figure 8 the Beta distribution maintains a finite lower bound at zero while infinitely extending to the right to capture the unlikely chance that a program more than doubles in cost. Estimators can modify the distribution shape through scaling to best represent a program's cost behavior. Kuhl and his colleagues (2009) developed a mechanism known as the Visual Interactive Beta Estimation System (VIBES) to manually fit beta distributions using sliding scales.

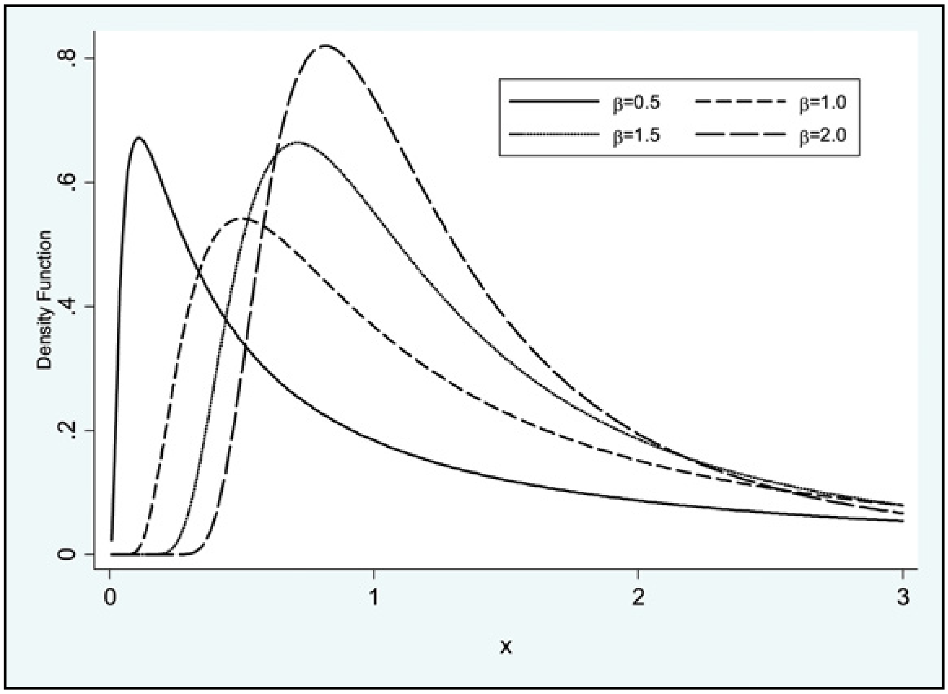


Figure 8. Probability Density Function of Beta Distribution. Source: Hanook et al. (2013).

In 2014, senior representatives from every branch of service and the National Aeronautics and Space Administration (NASA) convened to develop the *Joint Agency Cost Schedule Risk and Uncertainty Handbook* (CSRUH). The intent was to promulgate the best practices for establishing a "systematic, structured, repeatable and defendable process for delivering comprehensive estimates to Government leadership to get the best possible capability with increasingly limited available resources" (DOD and NASA 2014, ii). Experts included the Deputies Assistant to the Departments of Army, Air Force, Navy, NASA, and Christian Smart from the Missile Defense Agency. After studying every relevant distribution associated with cost growth, they concluded that the lognormal distribution (Figure 9) dominated all others in terms of frequency used for cost estimating.

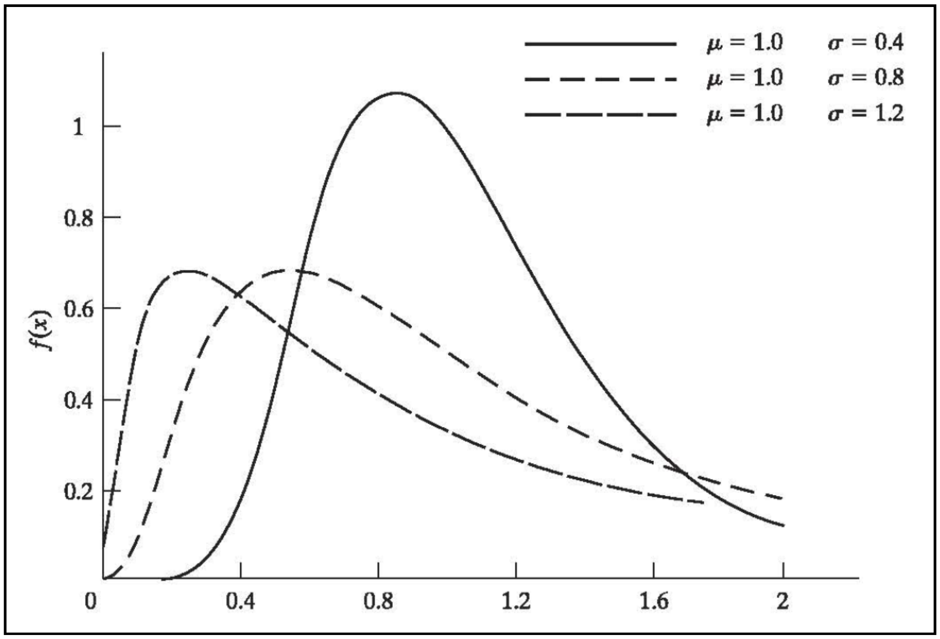


Figure 9. Lognormal Distribution Probability Density Functions. Source: Blanchard and Fabrycky (2011).

Nonetheless, the most appropriate distribution depends upon the application and the data that is available. Rather than choose one method, this thesis will consider all distributions relevant to uncertainty and decide based on goodness-of-fit tests. To promote consistency of program estimates, the CSRUH outlines the most effective uncertainty distributions and their typical applications and parameters in Table 2.

Table 2. CSRUH Recommended Uncertainty Distributions.
Source: DOD and NASA (2014).

DISTRIBUTION	TYPICAL APPLICATION	KNOWLEDGE OF MODE	NUMBER OF PARAMETERS REQUIRED	RECOMMENDED PARAMETERS
Lognormal	**Default** when no better info. Probability skewed right. Replicate another model result. Power OLS CER uncertainty.	Mean or median known better than the mode	2	**Median, high** (some tools have a 3rd parameter : "Location". By default, it is zero. Used to "shift" the lognormal left or right (even into negative region))
Log-t	Log-t when < 30 data points		3	**Add Degrees of Freedom**
Triangular	Expert opinion. Finite min/max. Probability reduces towards endpoints. Skew possible. Labor rates, labor rate adjustments, factor methods	Good idea	3	**Low, mode, and high**
BetaPert	Like triangular, but mode is 4 times more important than min or max.	Very good idea	3	**Low, mode, and high**
Beta	Like triangular, but min/max region known better than mode.	Not sure	4	**Min, low, high, and max**
Normal	Equal chance low/high. Unbounded in either direction Linear OLS CER uncertainty.	Good idea, but unbounded in either direction	2	**Mean/Median/Mode and high value**
Student's-t	t when < 30 data points		3	**Add Degrees of Freedom**
Uniform	Equal chance over uncertainty range. Finite min/max.	No idea	2	**Low and High** (some tools require min and max)
Empirical Fit	Unable to fit a distribution to the data	Not required	N/A	**Enter source data and estimated probability for each data point**
Note: *Low/high* are defined with an associated percentile *Min/Max* are the absolute lower/upper bound (also known as the 0/100)				

When determining the goodness-of-fit, the CSRUH recommends using the following: Kolmogorov-Smirnov (K-S), Anderson-Darling (A-D), and Chi-Squared (Chi^2) (DOD and NASA 2014). Furthermore, when comparing the fitted distributions, the CSRUH notes the value in employing the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) to compare the relative goodness-of-fit levels for each distribution. By combining quantifiable measures with recent historical data, this thesis provides just what the CSRUH recommends: realistic and objective cost estimates that capture uncertainty in a way that informs the decision maker (DOD and NASA 2014).

2. Cost Estimation

Professors Mislick and Nussbaum serve as NPS faculty and specialize in life-cycle cost estimating and modeling for government programs. In their book titled *Cost Estimation: Methods and Tools*, they provide a structured approach to cost estimation (CE) in what seems to be a very ubiquitous environment. Chapter 4 ("Data Sources") led this study's data collection efforts to the Cost Assessment Data Enterprise (CADE) to find the most relevant and exhaustive programmatic data; Chapter 5 ("Data Normalization") provides the standard on how to prepare the sponsor-provided data for analysis regarding inflation; Chapter 16 ("Cost Benefit Analysis and Risk Uncertainty") introduces a powerful risk tool/software but does not cover its employment in detail. As such, this study explores alternate software platforms that are more widely available to the DOD community. In this case, Microsoft Excel will serve as the primary platform and the Risk Simulator add-in from Real Options Valuation, Inc. acts as a means for building the visual products.

3. Excel-based Monte Carlo Simulations

Mun is the founder, chairman, and CEO of Real Options Valuation, Inc. (ROV). He is the creator of several powerful software tools used to manage risk (Mun 2015). Mun's book titled *Readings in Certified Quantitative Risk Management* provides detailed instructions on how to use the Risk Simulator Excel add-in for running Monte Carlo simulations. The software associated with Mun's book provides a powerful tool for leveraging ready-built probability distributions when modeling cost and schedule risk. Additionally, Risk Simulator provides the powerful and interactive visual products found throughout Chapter III and IV of this report.

D. THESIS FOCUS: PROGRAMMATIC RISK ANALYSIS

As history proves, cost overrun in defense spending is nearly inescapable; therefore, risk management is a necessity. Unfortunately, the previous results of cost-risk analysis provide perishable insights in the world of evolving technology and organizations. Most of the relevant prior work on the subject is nearing a decade old or too specific to apply to strategic-level decisions within AMA. Thankfully, many of their techniques are repeatable and compatible for building a novel methodology for capturing risk. To achieve that end, this thesis analyzes more than just individual programs, but also the overall portfolio risk related to TRAC-provided scenarios. The holistic measure of portfolio risk also serves as a metric for comparing the assessed operational impacts of selecting programs. The result is a quantifiable recommended cash reserve that supports the decision maker's desired confidence level for maintaining cost and schedule objectives without sacrificing technical performance or operational effectiveness. Finally, the results include program comparisons to identify the elements of the portfolios that contribute the most risk and thus aide in the prioritization effort. By mitigating subjectivity through data-driven distributions and improving foresight via Monte Carlo simulations, this thesis bridges the gap between specific program uncertainty and industry trends to develop an objective CE methodology that adequately informs investment decisions.

III. METHODOLOGY AND DATA PRESENTATION

To achieve the research objectives outlined in Chapter I, this thesis follows a five-step methodology: (1) data mining, (2) normalization, (3) distribution construction, (4) simulation, and (5) analysis and context of cost positions.

Figure 10 demonstrates the process by which the methodology follows for achieving the desired outcome of generating new cost variance benchmarks for acquisition programs based on commodity and milestone.

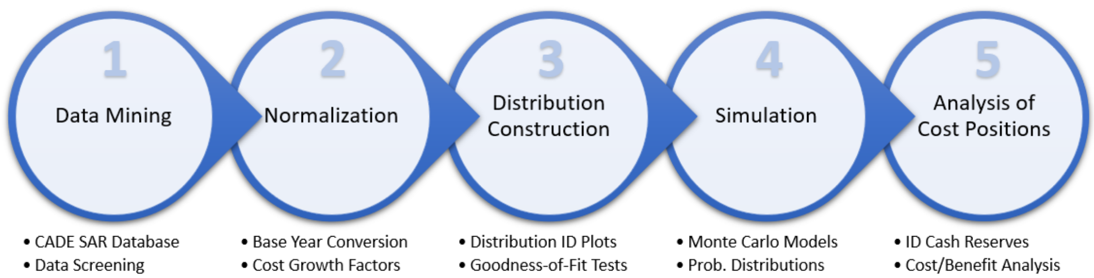


Figure 10. Thesis Methodology Process Flowchart

A. DATA MINING AND SCREENING

The process of data mining isolates the appropriate set of cost data that fits the conditions of this study. In this case, the research focuses on current and legacy U.S. Army acquisition programs that meet MDAP and ACAT-I requirements. Individual data points include each milestone report for prime and subcontractor programs while avoiding those that overlap.

This thesis relies on historical data from the Cost Assessment Data Enterprise (CADE) managed by the Office of the Secretary of Defense Cost Assessment and Program Evaluation (OSD CAPE). Before 2018, cost analysts had to scour dozens of individual service-managed databases to consolidate, organize, and analyze acquisition-related cost reports. With the launch of CADE's online cloud-based portal in December of 2018, authorized DOD officials can now query raw data, download detailed reports, submit contractor data, and perform rough analytics on a joint web-based application (CADE 2021). Military leaders in supervisory roles can grant DOD analysts access to CADE upon submission of a company-to-company nondisclosure agreement. Figure 11 illustrates the collective intent of CADE by consolidating cost platforms into an effective and easily searchable database.

Figure 11. Consolidation of Cost Data Platforms into CADE

The result is a significant decrease in the time spent on collecting and validating cost data. Therefore, estimators and researchers can quickly gather relevant and credible data from total program costs down to individual line-item estimates within the work breakdown structure (WBS) of major acquisition contracts.

The CADE portal hosts two main repositories of vetted reports while also providing crosstalk with endorsing organizations within each of the sister services. The Cost and Software Data Reporting (CSDR) interface presents detailed data reports on every major defense acquisition program (MDAP) and Acquisition Category (ACAT) I contract. Figure 12 depicts the browsing criteria and filtering functions built within the CSDR database. As seen in the bottom left of the figure, there are over 39,000 submissions and nearly 10,000 contractor cost data reports (CCDR) available for download.

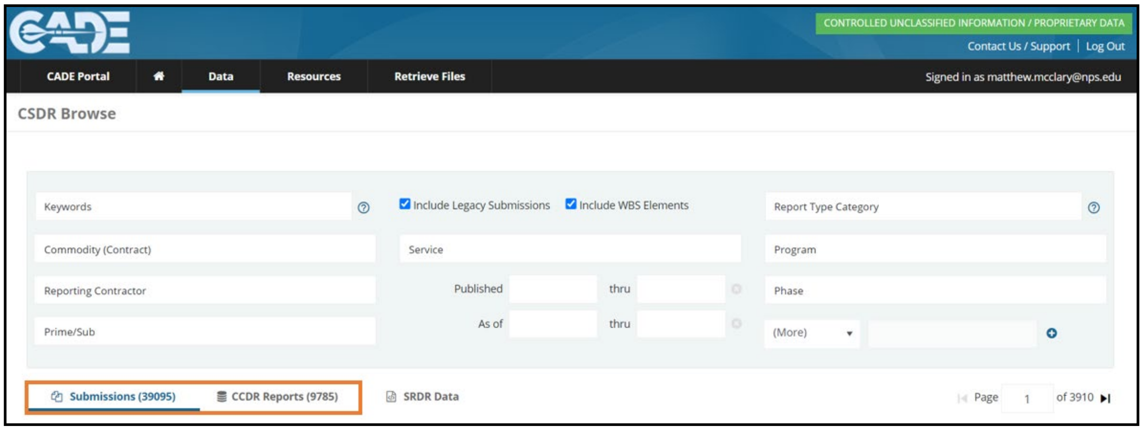


Figure 12. CADE Cost and Software Data Reporting Interface. Source: CADE 2021.

While the scrupulous nature of the CSDR database provides a mechanism for investigating individual programs, the Selected Acquisition Report (SAR) interface focuses on the comprehensive programmatic summary records and therefore provides the most effective means for collecting the data for this study. Figure 13 highlights the SAR interface for reviewing program level metrics organized by report type, service, and date range.

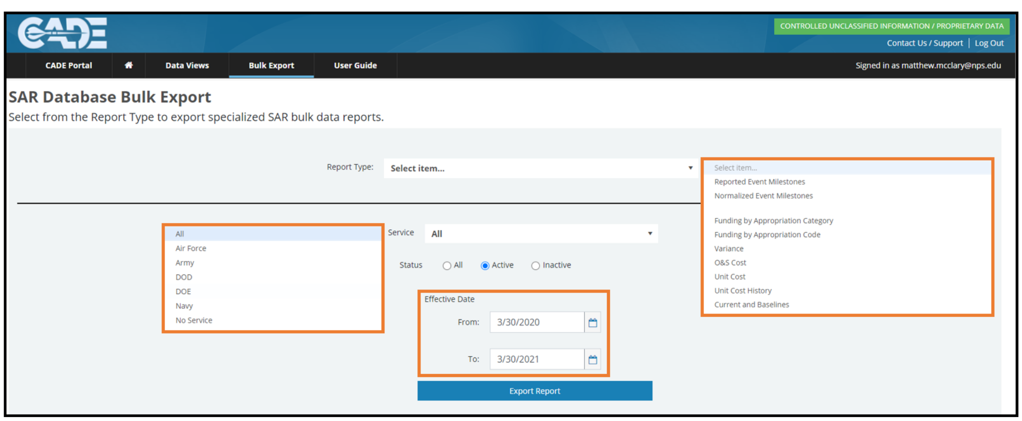


Figure 13. CADE Selected Acquisition Report (SAR) Interface. Source: CADE 2021.

The selected acquisition report provides the benefit of reviewing consolidated bulk data whereas the CSDR database separates all subcontracts from their prime program, thereby possibly introducing potential errors due to double counting. As noted in the far-right dropdown menu in Figure 13, the available report types include "current and baseline" estimates. By choosing this option, the output Excel file generates an organized table that presents available SAR data in "base year" dollars (BY$) depending on the start date of the contract.

Table 3. Current and Baseline Bulk SAR Data Snapshot. Adapted from NASA and DOD (2021).

Bulk Current and Baselines Funding
Database as of date: 3/8/2021
From: 3/29/1900 Active/Inactive: All

PNO	Service	MilStd	Program Type	Context Tags	Base Year	Current Total $	Baseline Total $	Total Cost Change	Cost Growth Factor
148	Army	Missile	MDAP	Milestone B SAR	1988	$ 5,746,800,000.00	$ 2,783,200,000.00	$ (2,963,600,000.00)	2.064817476
148	Army	Missile	MDAP	Milestone C SAR	1988	$ 2,285,600,000.00	$ 1,284,400,000.00	$ (1,001,200,000.00)	1.779507941
148	Army	Missile	MDAP	Latest Representative SAR	2002	$ 10,612,600,000.00	$ 5,505,800,000.00	$ (5,106,800,000.00)	1.927530967
179	Army	Helicopter	MDAP	Earliest SAR	2005	$ 3,149,100,000.00	$ 2,790,300,000.00	$ (358,800,000.00)	1.128588324
179	Army	Helicopter	MDAP	Milestone B SAR	2005	$ 3,159,500,000.00	$ 2,790,300,000.00	$ (369,200,000.00)	1.132315522
179	Army	Helicopter	MDAP	Latest Representative SAR	2005	$ 501,700,000.00	$ 2,790,300,000.00	$ 2,288,600,000.00	0.179801455
182	Army	Helicopter	MDAP	Earliest SAR	2006	$ 1,638,300,000.00	$ 1,635,100,000.00	$ (3,200,000.00)	1.001957067
182	Army	Helicopter	MDAP	Milestone C SAR	2006	$ 1,647,900,000.00	$ 1,635,100,000.00	$ (12,800,000.00)	1.007828267
182	Army	Helicopter	MDAP	Latest Representative SAR	2006	$ 1,626,900,000.00	$ 1,635,100,000.00	$ 8,200,000.00	0.994985016
202	Army	Helicopter	MDAP	Earliest SAR	2006	$ 6,553,000,000.00	$ 5,507,400,000.00	$ (1,045,600,000.00)	1.189853651
202	Army	Helicopter	MDAP	Milestone B SAR	2006	$ 7,117,800,000.00	$ 5,507,400,000.00	$ (1,610,400,000.00)	1.29240658
202	Army	Helicopter	MDAP	Milestone C SAR	2010	$ 10,452,500,000.00	$ 5,507,400,000.00	$ (4,945,100,000.00)	1.897901006
202	Army	Helicopter	MDAP	Latest Representative SAR	2010	$ 12,476,900,000.00	$ 8,856,900,000.00	$ (3,620,000,000.00)	1.408720884
205	Army	Missile	MDAP	EarliestSAR, Milestone B SAI	2009	$ 4,856,600,000.00	$ 3,316,000,000.00	$ (1,540,600,000.00)	1.464595899
205	Army	Missile	MDAP	Latest Representative SAR	2009	$ 6,297,100,000.00	$ 3,316,000,000.00	$ (2,981,100,000.00)	1.899004825

Using Excel's filter function, the analyst sorts the data to isolate one service (i.e., Army), remove redundant data points (annual/quarterly SAR reports between milestones), and filter out incomplete reports. Often, developmental programs do not include current estimates due to their infancy while some mature programs lack baseline estimates due to their smaller initial size. This is because only those contracts that meet the acquisition thresholds for ACAT I or MDAP must submit routine SAR data. The following list describes the five-step screening process for identifying the data that acts as the foundation of this study's analysis:

1. Remove items that lack baseline estimates by unchecking "(blanks)" in "Baseline Total $" column of the Excel spreadsheet. This is most often a result of smaller programs escalating to the ACAT-I or MDAP tier after their inception, thereby introducing a new requirement to submit SAR data. Programs that lack baseline data cannot return comparable cost growth factors without an initial cost estimate. This step reveals a full

report of data capable of CGF comparison and reduces the total number of reports from 4,451 to 2,158.

2. Remove interim and erroneous SAR reports to avoid redundancy and unintended weighting of longer duration programs by unchecking "(blanks)" in the "Context Tags" column of the Excel spreadsheet. CADE analysts add "Context Tags" to every line item upon entry of SAR data to indicate the reports status in relation to acquisition milestones (A, B, or C). In addition, CADE annotates whether the report is the earliest or latest SAR for that program. This step reveals only the major event reports associated with earliest/latest estimates and milestones A, B, and C in the acquisition life cycle thereby reducing the number of reports from 2,158 to 389.

3. Remove the "latest representative SAR" of terminated and transitioned programs to prevent unwarranted skewness in the data. This is a two-step process. First, one must filter out all programs other than those labeled as "terminated" in the "Status" column of the spreadsheet. Then, the analyst must filter by "context tag" to reveal the latest representative SAR" before manually removing or striking the remaining line items. This process repeats itself for programs classified as "transitioned/restructured." The premature ending of programs reflects an inadmissible cost savings since they do not deliver the intended operational benefit and therefore become an overall cost burden. This step reveals complete estimates for programs that are active, operational, transitioned, fully-developed or terminated, and thereby reduces the data pool from 389 to 377.

4. Scope the remaining data by commodity to generate a dataset that captures the three commodities of interest (C3I, Ground Vehicles, and Rotary Wing Aircraft). During this step, the analyst must scrub related commodity types to maximize sample size. For instance, the Joint Light Tactical Vehicle (JLTV) falls within the "Ground Combat" category but can certainly relate

to ground vehicle interest as well. In this case, each commodity of interest includes two listed categories:

- C3I = (“C3I”) + (“Aircraft C3I”) = 44 reports
- Rotary = (“Helicopter”) + (“Helicopter System”) = 41 reports
- Vehicle = (“Ground Vic”) + (“Transport Vic”) = 17 reports

5. Scope the remaining data by milestone to reveal the final subsets required for analysis. In the end, there are 102 relevant reports.

B. NORMALIZATION OF DATA

Before this study can build simulations and conduct statistical analysis, the process must account for variations across programs to achieve consistent and comparable metrics (Mislick and Nussbaum 2015). This thesis concentrates on cost growth factors (CGF) that are indifferent of base year (BY) or inflation by calculating comparable ratios for each respective data point. As seen in the shaded area on the right of Table 3 in the previous section, the “current total” divided by the “baseline total” reveals the respective CGF values.

Fortunately, the availability of consistent base year (BY) metrics between programs within the CADE data satisfies the need to normalize against inflation indices. Several cost estimating experts proved this point in the 2012 International Cost Estimating and Analysis Association (ICEAA) symposium as part of their effort to update S-curve CE benchmarks using CGF from SAR data (Lee et al. 2012). Figure 14 outlines the simplicity of using BY estimates considering the ineffectual nature of inflation indices.

CGF is invariant with Base Year

– Addition of BY12 should not necessitate new CGFs, since the CGFs based on BY12 would be mathematically identical to those based on BY

$$CGF\ (BY) = \frac{CE\ (BY)}{BE\ (BY)} = \frac{CE\ (BY) \cdot i}{BE\ (BY) \cdot i} = \frac{CE\ (BY12)}{BE\ (BY12)} = CGF\ (BY12)$$

where *CE* = Current Estimate, *BE* = Baseline Estimate, *BY* = Base Year, *i* = escalation index from BY to BY12

Figure 14. Proof of Invariance of Cost Growth Using Base Year Estimates. Source: Lee et al. 2012.

To ensure consistency and accuracy across four decades of historical records, CADE analysts add “context tags” (column five of Table 3) to each line item to ensure uniformity between reports based on changes to programmatic reporting requirements. For instance, reports submitted prior to 2001 followed numerical milestones while newer reports reflect the alphabetical milestones familiar to today’s acquisition community. This effort greatly improves the ability for this study to analyze variance between milestones considering the tradeoff between uncertainty and known costs as the program matures throughout the system life cycle. Figure15 illustrates the ever-changing methods of cost estimation as programs evolve overlaid with the phases of technology maturation running along the top of the illustration. Along the bottom of the figure, one can see the decreasing uncertainty as estimate fidelity improves and contractors report their actual expenses. Chapter IV of this thesis will analyze the differences in cost variance between milestones.

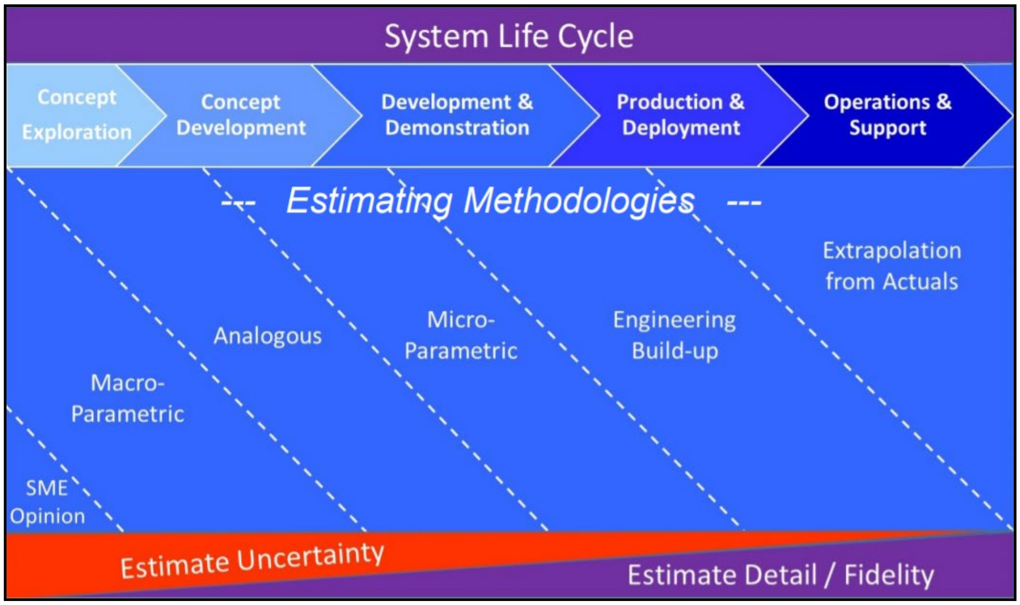


Figure 15. Estimating Methodologies Overlaid with System Life Cycle. Source: Alexander (2020).

Initial review of the raw CGF values for all Army programs reveals a right-skewed distribution seen in Figure 16. As recommended by the CSRUH, this study leverages the Mann-Wald theorem for determining the appropriate bin size. In this case, the method dictates 24 bins ranging from 0.5 to 12. In this case, the minimum CGF for a program that achieve full operational capability was the first increment of small, unmanned air and ground sensors for the Early Infantry Brigade Combat Team (E-IBCT) at -52%, while the highest CGF belongs to the Apache Longbow Helicopter at a staggering 1,058% cost growth. Chapter IV of this thesis will explore the frequency and probability of experiencing dramatic cost growth depicted by the outliers annotated in Figure 16. The obvious right-skewness and dramatically low kurtosis in the figure indicates the need for transforming the data before conducting statistical analysis.

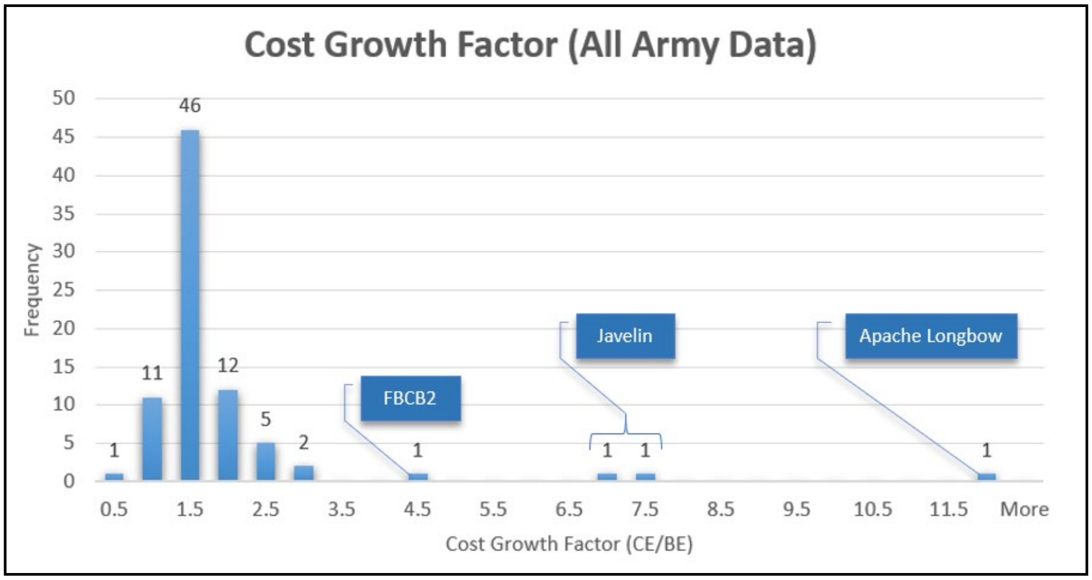


Figure 16. Cost Growth Factor Histogram for All Army Programs

Figure 17 shows the new frequency distribution after transforming the data logarithmically. The resulting histogram reveals potential normality. The initial transformation acts as an exploratory test for determining the need to conduct distribution identification plots via ready-built software statistics packages. This study relies on Minitab to test all relevant transformations; Minitab provides an expedited and relatively inexpensive platform for conducting statistical analysis.

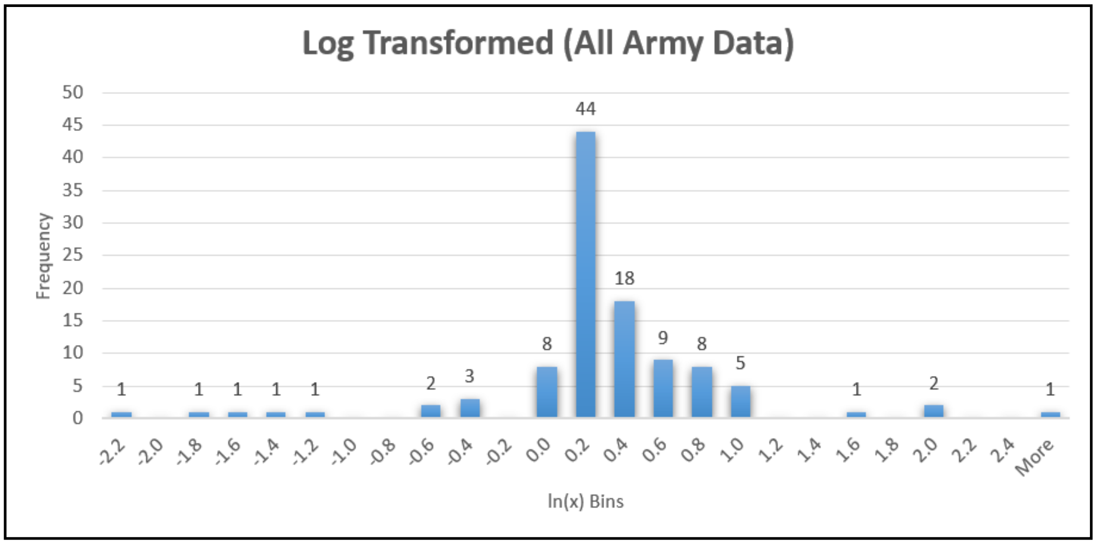


Figure 17. Log Transformed CGF Histogram for All Army Programs

C. DISTRIBUTION CONSTRUCTION

Identifying the most appropriate probability distribution for predicting cost growth behavior is a critical step for establishing credible results. Instead of applying a catch-all distribution that ignores the data-driven approach that this study embraces, this process leverages historical CADE data and powerful statistics software packages to identify and construct relevant distributions.

After calculating the CGF for every program at each milestone, the resulting data undergoes a "Distribution ID Plot" using Minitab to determine the best fitting probability distribution (Figure 18). This process quickly compares all relevant distributions and applies a goodness-of-fit test to determine the best fit.

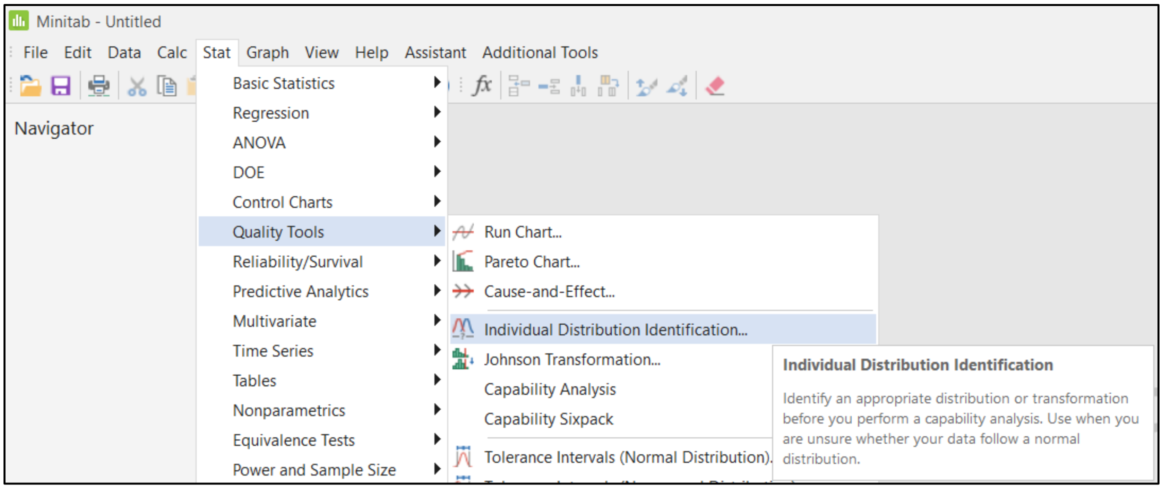


Figure 18. Minitab Distribution ID Plot Function.[1]

When considering all available distributions, this study limits the feasible alternatives to those recommended by the CSRUH and highlighted in Table 2 (Lognormal, Triangular, BetaPert, Beta, Normal, Uniform, and Empirical Fit). Considering Christian Smart's caution that "the devil is in the tails" (2020, 171), the Anderson-Darling score for each distribution presents the ideal metric for assessing goodness-of-fit since it captures the most accuracy along the tails (DOD and NASA 2014). Alternatively, the Kolmogorov-Smirov test focuses on the center of the distribution while the Chi-Squared offers the greatest ease of use, but also includes a sensitivity to the number of bins used to stratify the data (DOD and NASA 2014). This process repeats for each commodity and milestone report to determine the appropriate distribution parameters for each factor before building a cost simulation. Figure 19 includes an example Minitab output for determining the distribution parameters outlined on the shaded right-side table. These values provide the building blocks for constructing relevant distributions and achieving thesis objective #1.

[1] Portions of information contained in this publication/book are printed with permission of Minitab, LLC. All such material remains the exclusive property and copyright of Minitab, LLC. All rights reserved.

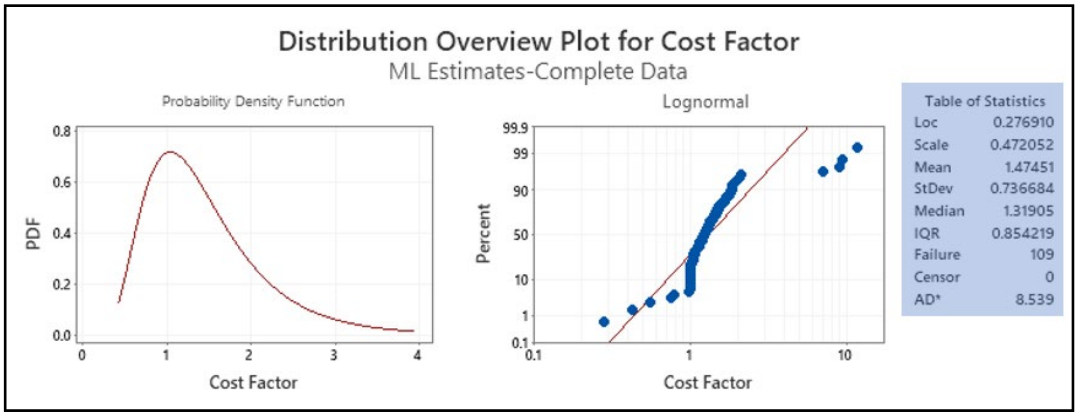


Figure 19. Minitab Distribution Overview Pot and Parameters

D. SIMULATION

According to the CSRUH, using custom-based simulation data to generate insights into cost uncertainty is the ideal method (2014) considering its ability to run thousands of stochastic models in mere seconds. After identifying the commodity and nearest milestone of each program within the TRAC portfolios, this study builds bespoke simulations to assess cost variance.

Using Excel's Risk Simulator add-in, this thesis captures the uncertainty associated with total program cost by leveraging Monte Carlo simulations influenced by the tailored distributions. Consistent with cost estimating industry standards, the simulation profile includes 10,000 iterations and generates a probability plot that mirrors Figure 20 (DOD and NASA 2014, 54). The top-left portion of Figure 20 provides a mechanism for rapidly calculating one-tail and two-tail critical values associated with the DM's certainty level; this procedure satisfies thesis objective #2.

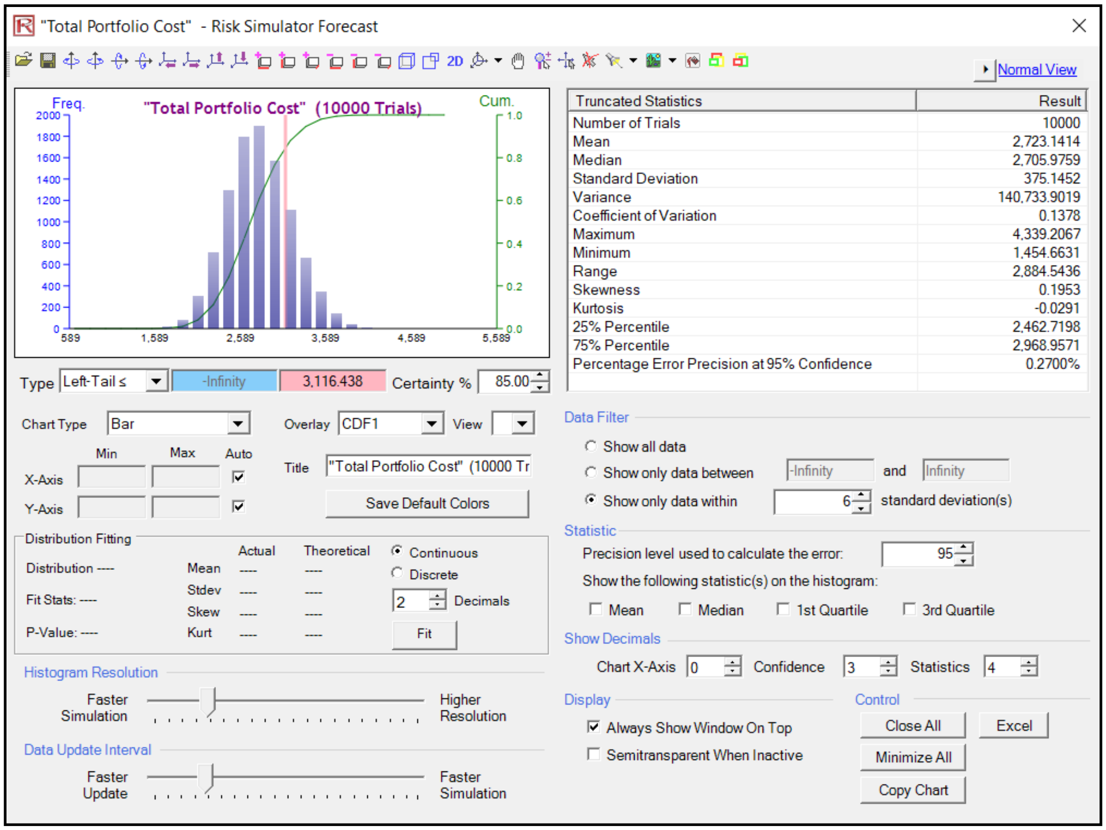


Figure 20. Risk Simulator Output Example

Lastly, using a subset of the historical data from CADE, this research employs a cross-validation technique to test the improved methodology by running the simulation against real-world baseline estimates reserved for model testing. An analysis of variance (ANOVA) test between actual current estimates and modeled results will serve as the mechanism for validating the methodology.

E. ANALYSIS OF COST POSITIONS

Statistical inferences from the cost variance models allow this study to assess the appropriate level of contingency cash reserves associated with TRAC-defined cost portfolio positions to inform program prioritization and operational effectiveness. The probability plot of total cost generated by Risk Simulator aides in determining the necessary cash contingency reserve level based on the decision maker’s confidence level. According to Smart, the recommended funding level should reside above 80% to capture the “exception variation” found within the tails of lognormal distributions (2020). The difference between the modeled total cost at the dictated confidence level and the TRAC-provided point estimates reveals the cash reserve level and represents the financial risk associated with that scenario. This process achieves thesis objective #3, which is to offer data-driven information during program selection.

By comparing the modeled cost for each funding profile with the assessed operational effectiveness benefit, this thesis provides a cost/benefit analysis based on the total program cost acting as the predictor to facilitate an informed decision on which alternative is best (thesis objective #4). In addition, this research provides updated CGF benchmarks by commodity to aide in future cost estimating efforts.

IV. RESULTS OF DATA ANALYSIS

A. DATA MINING AND REFINEMENT

Using the "bulk" data query from CADE's SAR database, this study collected estimates on every ACAT-I or MDAP program since 1973. The resulting spreadsheet yielded 4,480 data points between every service of the DOD. To ensure accuracy of analysis, this research screened the data using Excel's filter function applied in accordance with the criteria outlined in Figure 21 and described in this section. Figure 21 illustrates the screening process described in Chapter III and the resulting quantity of nonredundant and relevant reports for all DOD programs.

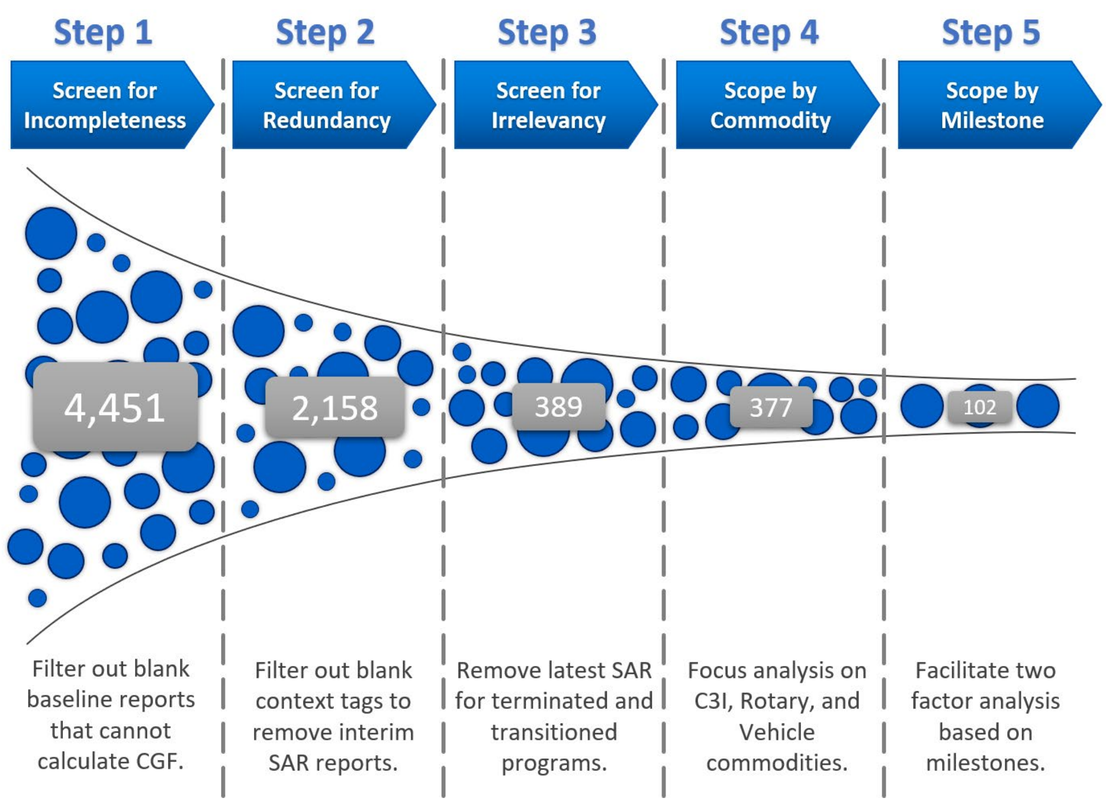


Figure 21. Screening Process for CADE Data Refinement

Although the number of relevant data reduces dramatically, the net result still provides adequate datasets for performing statistical analysis. Table 4 provides the number of data points by category that will participate in the upcoming ANOVA tests.

Table 4. Number of Data Points by Commodity and Milestone Matrix

Commodity	**Earliest**	**MS-A**	**MS-B**	**MS-C**	**Latest**	**Total**
C3I	10	0	14	5	15	44
Rotary Wing	7	0	10	9	15	41
Vehicles	2	0	6	2	7	17
Total	19	0	30	16	37	102

According to the author of *Cost Estimation: Methods and Tools*, the absence of Milestone-A reports is not uncommon since many programs forgo the first milestone event when operational necessity dictates their importance (Gregory Mislick, personal communication, April 23, 2021). However, for the purpose of this study, the earliest SAR data points will act as Milestone-A reports since they occur in the first phase of the acquisition life cycle.

B. NORMALIZATION OF DATA AND ANOVA

After calculating the CGF for each SAR report, the resulting metric provides a comparable and unitless measurement of cost overrun or savings so long as the analyst ensures that the current and baseline estimates reflect the same base year dollars. In every case, CADE analysts converted all SAR data for each specific program into a common base year currency to avoid inflation-induced errors. For this study, all costs have been baselined to the program's origination year to facilitate CGF calculation. After confirming common base years, the CGF values migrate from Excel to Minitab for ease of statistical analysis. This study employs the following analysis techniques to determine the nature of the data before answering the key research questions outlined in Chapter I:

1. Analyze the overall trend of variance in CGF regardless of commodity or milestone using a one-way ANOVA test between Army data and the entire DOD/DOE database.

2. Analyze the trend of variance in CGF between commodities using a one-way ANOVA between C3I, Rotary, and Vehicle data.

3. Analyze the trend of variance between CGF milestones using a one-way ANOVA between Milestones A, B, C, and the Latest SAR data.

4. Analyze the trend of variance in CGF within each commodity and between milestones using a two-way ANOVA test.

1. Trend of Variance between Army-DOD

In this test, the null hypothesis posits that the mean CGF for all Army data is equal to the mean CGF of all DOD/DOE data, thereby suggesting that benchmark CGFs would be indifferent to their service. Figure 22 summarizes the overall trend of variance within the Army data (N = 97) relative to the entire DOD enterprise and the Department of Energy (DOE) (N = 377). Considering that the resulting p-value (0.957) is far greater than the significance level ($\alpha = 0.05$), this test concludes that there is no significance statistical difference between the Army's mean CGF value and the entire DOD/DOE. As seen in the relatively even interval plot in Figure 22, the Army CGF data demonstrates a balanced comparison with the DOD/DOE. Therefore, future analysis includes sister service data to promote adequate sample sizes. The increased sample size will aide in the robustness and accuracy of findings when modeling the variance between commodities and milestones for predicting cost growth. In essence, regardless of military service, the benchmark value for CGF appears to fall at approximately 1.55 or 55% cost overrun across all DOD acquisition programs.

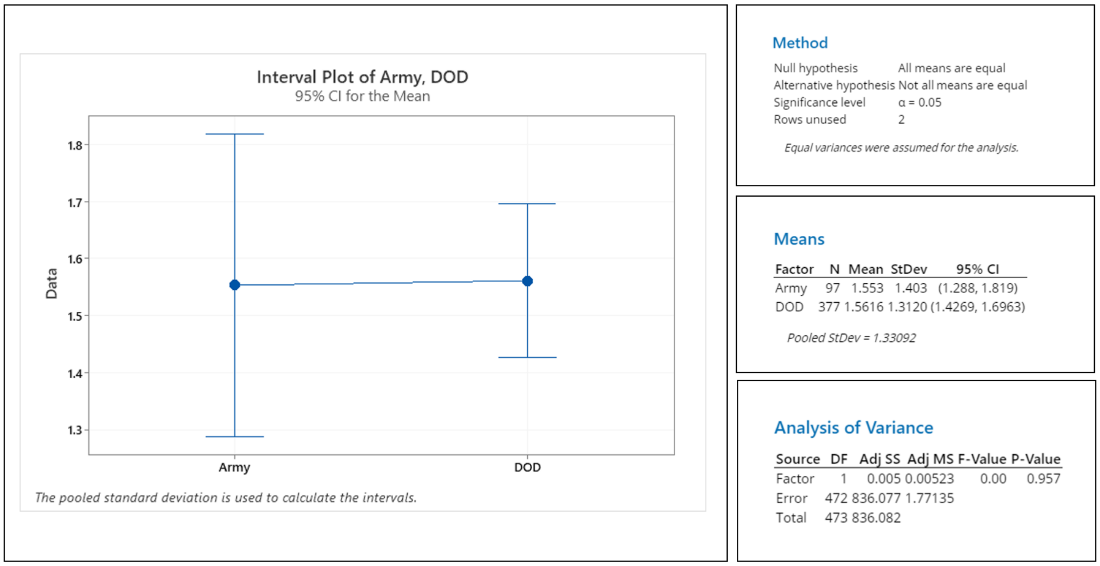
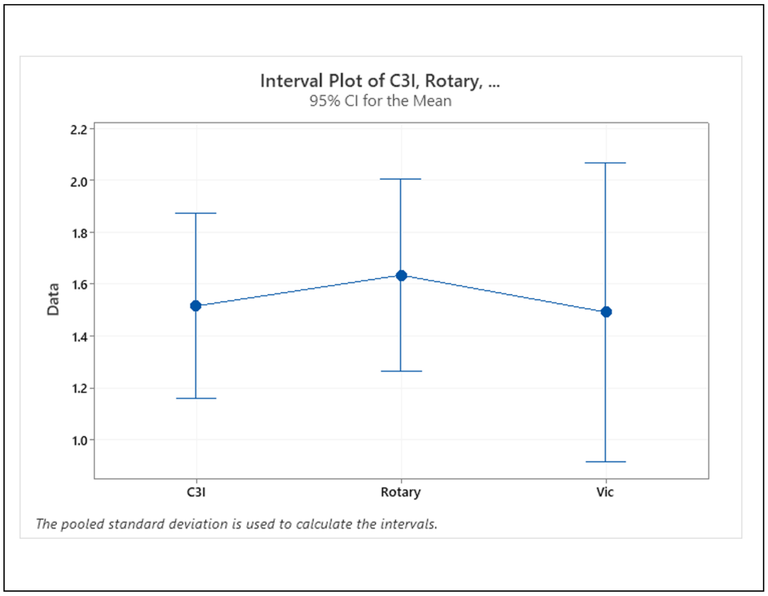


Figure 22. Minitab One-way ANOVA Test Results for Army-DOD

2. Trend of Variance between Commodities

In this test, the null hypothesis postulates that the mean CGF for all three commodities are equal, thereby suggesting that benchmark CGFs would be indifferent to their technology type or commodity classification. Figure 23 reveals the overall trend of variance within each commodity relative to one another. In this case, the test fails to reject the null hypothesis that all means are equal based on the p-value (0.872) exceeding the significance level ($\alpha = 0.05$). All three commodity types do not appear to have a statistically significant different mean CGF values as seen in the balanced interval plot in Figure 23. In short, the benchmark values for all three commodities span a similar 1.55 CGF mark seen in the DOD/DOE data pool. The overlapping CGF values seen between commodities support the application of a single benchmark toward all ACAT-I and MDAP programs regardless of commodity or service.



Method

Null hypothesis	All means are equal
Alternative hypothesis	Not all means are equal
Significance level	α = 0.05

Equal variances were assumed for the analysis.

Means

Factor	N	Mean	StDev	95% CI
C3I	44	1.516	0.757	(1.159, 1.873)
Rotary	41	1.634	1.633	(1.265, 2.004)
Vic	17	1.492	0.778	(0.918, 2.067)

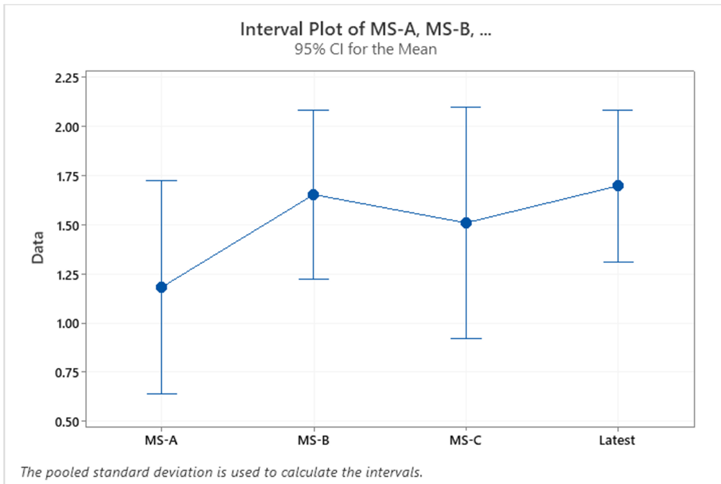
Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Factor	2	0.390	0.1951	0.14	0.872
Error	99	141.036	1.4246		
Total	101	141.426			

Figure 23. Minitab One-way ANOVA Test Results for C3I-Rotary-Vehicle

3. Trend of Variance between Milestones

In this test, the null hypothesis states that the mean CGF of each milestone is equal, thereby suggesting that benchmark CGFs would be the same regardless of their relative position in the acquisition life cycle. Figure 24 uncovers the overall trend of variance between milestones. The test fails to reject the null hypothesis that all means are equal based on the p-value (0.454) exceeding the significance level (α = 0.05). That being so, all four milestone report types bear no statistically significant difference in CGF. Although the interval plot seen in Figure 24 depicts a slight upward trend across milestones, the overlapping nature of the confidence intervals could just as easily reveal a negative sloping trend should the population means differ from the test results. Essentially, the benchmark values for each milestone once again span the equivalent 1.55 CGF mark seen in the DOD/DOE and commodity datasets, but the statistical significance is diminishing as reflected in the decreasing p-value across the three one-way ANOVA tests. As such, the value of this analysis lies within the variance or standard deviation captured within each of the factors tested.



Method

Null hypothesis	All means are equal
Alternative hypothesis	Not all means are equal
Significance level	α = 0.05

Equal variances were assumed for the analysis.

Means

Factor	N	Mean	StDev	95% CI
MS-A	19	1.1830	0.2328	(0.6433, 1.7227)
MS-B	30	1.654	1.900	(1.224, 2.083)
MS-C	16	1.510	0.448	(0.922, 2.098)
Latest	37	1.698	0.899	(1.311, 2.085)

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Factor	3	3.710	1.237	0.88	0.454
Error	98	137.716	1.405		
Total	101	141.426			

Figure 24. Minitab One-way ANOVA Test Results for Milestones A, B, C, and Latest

4. Trend of Variance by Commodity and Milestones

Before conducting the two-way ANOVA test, Excel's pivot table function facilitated the production of a three-dimensional surface plot of CGF averages. Figure 25 illustrates the average cost growth within each factor relative to the others. The peaks and valleys of the graph articulate the high and low CGF values under those specific conditions. As seen, the highest average CGF occurs within the rotary commodity at milestone B, while the C3I and vehicle areas experience higher cost growth in their latest SAR report (arrows indicate high points). The extremes of the surface plot provide insights in potential interactions between factors while the two-way ANOVA tests for statistical significance.

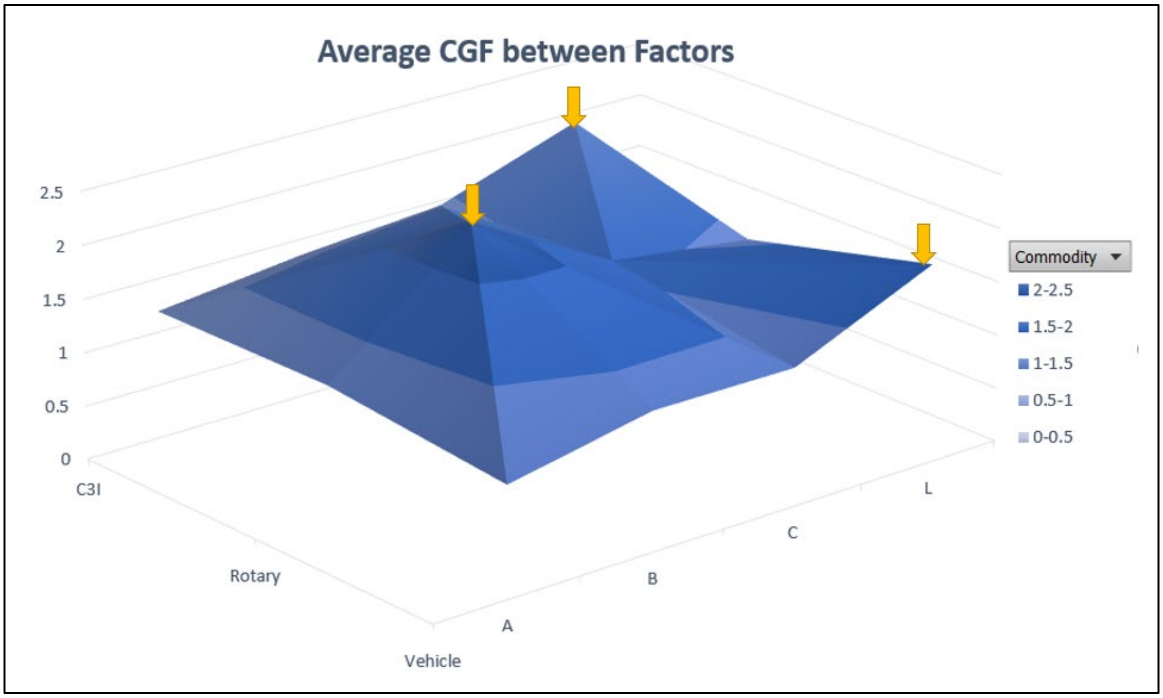


Figure 25. Surface Plot of Average CGF between Factors

Finally, the two-factor ANOVA test expresses the statistical significance of variance trends between commodity and milestone events. Figure 26 outlines the results of the test. In this case, the null hypotheses propose that (1) the means of all commodity groups are equal, (2) that the means of all milestone subsets are equal, and (3) that there is no interaction between commodity and milestone that effect the main CGF metric. Judging from the Minitab output in Figure 26, the two-way ANOVA test fails to reject all three null hypotheses considering that the p-values for each factor and their interaction exceed the significance level. In other words, commodity, milestone, and their interaction are not statistically significant factors in determining a program's overall CGF. Furthermore, the linear model created as part of the two-way ANOVA only explains a mere 8.55% (R-squared) of variation in a program's CGF. Thus, one cannot use this model to infer statistical relationships between variables outside of the sample data. Once again, the overall CGF remains unaffected by commodity, milestone, and the interaction between the two factors. Considering that there are no valid interactions between commodity and milestone, this study presents two independent cost simulations for comparison to determine the ideal method for capturing cost variance within a specific portfolio.

Factor Information

Factor	Type	Levels	Values
Commodity	Fixed	3	C3I, Rotary, Vehicle
Milestone	Fixed	4	A, B, C, L

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
1.19879	8.55%	0.00%	0.00%

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Commodity	2	0.799	0.3993	0.28	0.758
Milestone	3	3.298	1.0995	0.77	0.517
Commodity*Milestone	6	7.893	1.3156	0.92	0.488
Error	90	129.339	1.4371		
Total	101	141.426			

Figure 26. Minitab Two-way ANOVA Test Results for Commodity and Milestone

5. Overall Trend of Variance

After analyzing the variance trends within and between the commodity and milestone factors, this study concludes that there is no significant difference between mean CGF values. However, considering that averages do not capture the variation in cost data that contributes to risk, the benefit of this study is the measurement of variance within each of the categories themselves. Figure 27 illustrates the common thread through every tested factor based on the average CGF, thereby confirming the similarity between means. The within sample variances indicate areas that modeling and simulation techniques may prove useful in gaining insights. To compare the factors based on variation, this study leverages the coefficient of variation (CV) statistic recommended by the CSRUH (DOD and NASA 2014, 57).

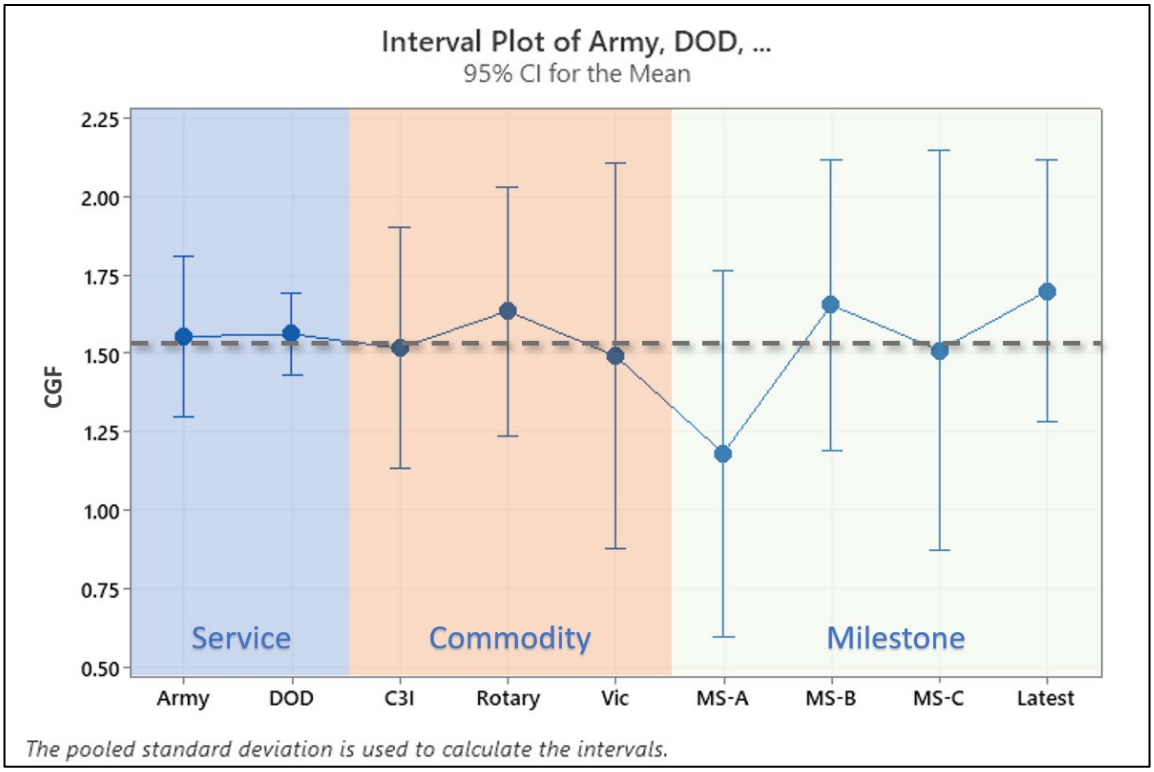


Figure 27. Adapted Minitab Interval Plot of Mean Confidence Intervals for All Factors

Table 5 illustrates the CV values associated with each variable. The CV values provide a normalized metric for capturing the dispersion within each factor (DOD and NASA 2014). Equation 1 demonstrates how to calculate CV values in a data set.

$$CV = StDev/Mean \tag{1}$$

The rotary commodity and Milestone-B factors harbor the greatest variation with CV values of 100% or more (rows shaded in Table 5). In other words, programs involving rotary technology nearest milestone-B carry the greatest potential for cost overruns considering the relatively equal CGF shared between all variables. This finding corroborates the insights gleaned from the 3D surface plot in Figure 25. In the next section, this study applies variance trends toward building bespoke distributions that model a program's cost variance.

Table 5. Coefficients of Variation for Cost Growth Factors

Factor	Mean	StDev	CV
C3I	1.516	0.757	50%
Rotary	1.634	1.633	100%
Vehicle	1.492	0.778	52%
MS-A	1.183	0.233	20%
MS-B	1.654	1.900	115%
MS-C	1.510	0.448	30%
MS-L	1.698	0.899	53%

C. DISTRIBUTION CONSTRUCTION

Although commodity type and milestone do not generate a statistically significant impact on the mean CGF for any given program, their respective variances influenced the overall cost risk associated with specific program portfolios. As such, this thesis identifies the appropriate distribution for each variable via Minitab's "Distribution ID Plot" function. During this process, the software attempts to fit the data to distributions recommended by the CSRUH, which include the Beta, normal, Weibull, lognormal, and three-parameter lognormal (TPLN) distributions. In this case, the TPLN does not apply since all CGF values are non-negative (DOD and NASA 2014) and Minitab does not test for the Beta distribution. Therefore, this research included the exponential distribution as an alternative to the Beta. In summary, this study tested the goodness-of-fit for each factor using the following distributions: exponential, normal, Weibull, and lognormal.

1. Distribution Identification Plots

After specifying the relevant distributions for testing, Minitab generated a probability plot and goodness-of-fit test statistics using the Anderson-Darling (AD) technique to produce a graphic like Figure 28. Since the AD formula measures the total area between the sample and the fitted cumulative density function (CDF), a lower score indicates a better fit (DOD and NASA 2014). This research chose the AD goodness-of-fit test based on its unique trait for capturing accuracy within the tails and compatibility with small sample sizes (Romeu 2003). The shaded areas on the right of Figure 28 highlight the goodness-of-fit graph and AD metric for the lognormal. The near linear trend seen in the probability plot indicates a normal distribution of the transformed data that can provide statistical inference.

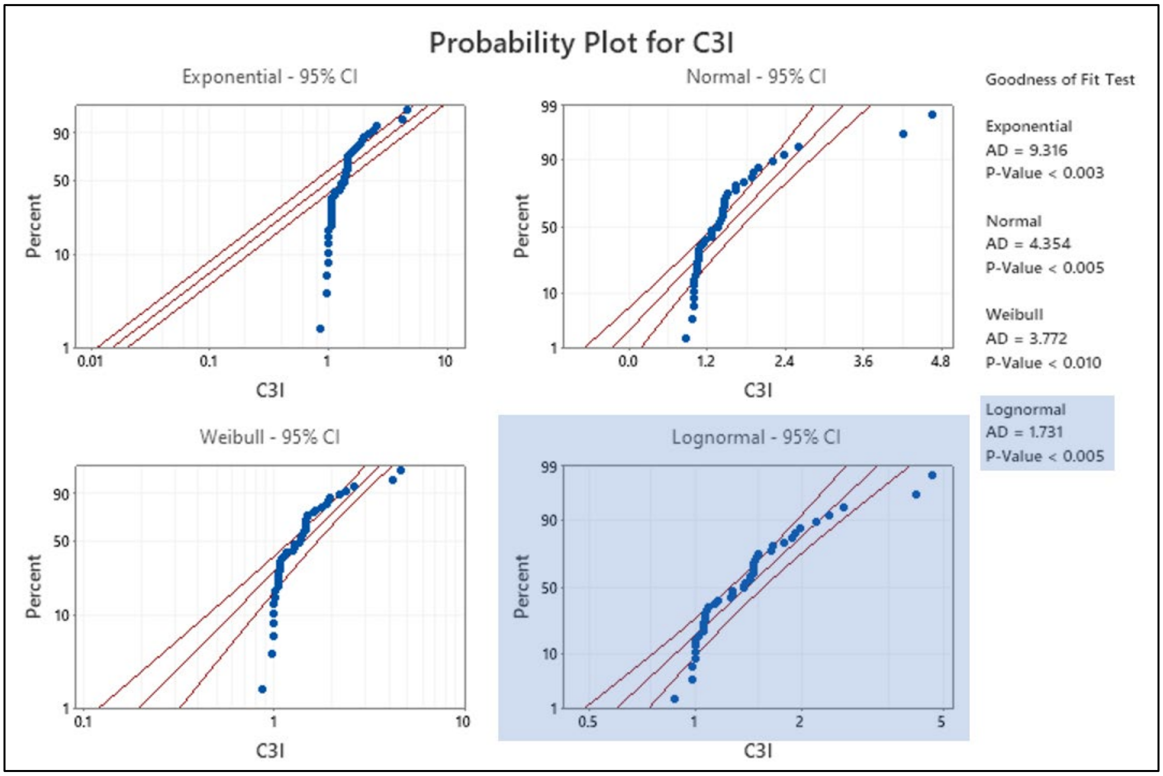


Figure 28. Minitab Probability Plot and Goodness-of-Fit for C3I

As seen in the shaded right column of Table 6, the lognormal distribution yielded the lowest —and therefore the best—AD score for every factor by achieving the lowest AD score when compared to the three other alternatives (reference Appendix A for all distribution identification plots). Summing the total AD score for each distribution reveals obvious choice and dominance of the lognormal vice the exponential, normal, and Weibull options.

Table 6. Anderson-Darling Scores for Goodness-of-Fit Test

	Anderson-Darling Scores			
Factor	**Exponential**	**Normal**	**Weibull**	**Lognormal**
C3I	9.316	4.354	3.772	1.731
Rotary	8.943	9.145	6.663	2.643
Vehicle	3.611	2.040	1.769	1.279
MS-A	6.069	1.330	1.363	1.154
MS-B	6.519	7.617	5.534	2.876
MS-C	3.966	0.617	0.634	0.346
MS-L	6.775	3.558	2.778	1.359
Total	**45.199**	**28.661**	**22.513**	**11.388**

2. Individual Distribution Overviews and Parameters

The next step of the process includes harvesting individual distribution parameters for the purpose of modeling the variance associated within each of the identified factors. The Minitab "Distribution Overview Plot" provides a mechanism for generating the mean and standard deviation parameters that accompany each individual lognormal distribution. Figure 29 shows the truncated output (hazard and survivability plots excluded due to irrelevance) of the distribution overview plot and parameters for C3I as an example for each (see reference Appendix B for the entirety of plots for every factor). In every case, the transformed data follow a near linear relationship when plotted along the lognormal graph as seen on the right graph of Figure 29.

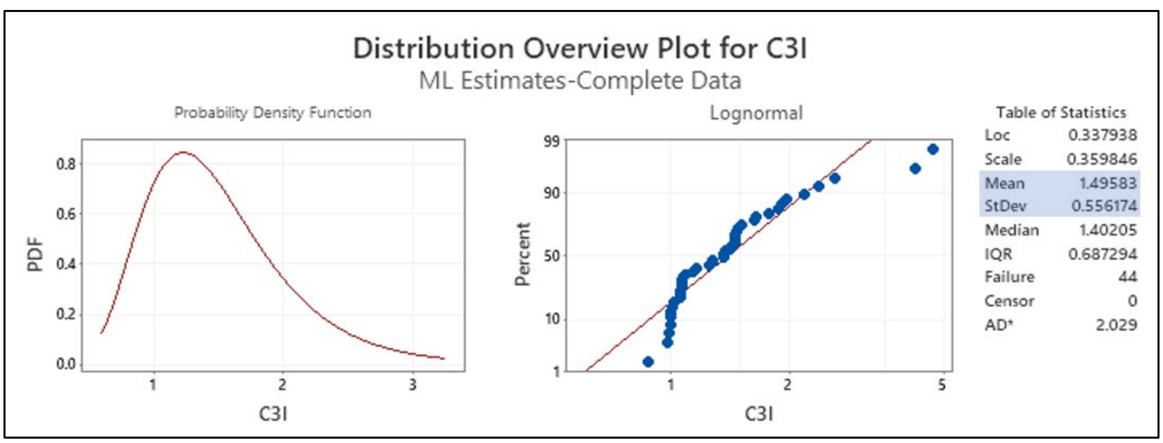


Figure 29. Minitab Distribution Overview Plot and Parameters for C3I

Table 7. Lognormal Distribution Parameters

Lognormal Parameters		
Factor	**Mean**	**StDev**
C3I	1.4958	0.5562
Rotary	1.5401	0.6571
Vehicle	1.4707	0.5677
MS-A	1.1821	0.2123
MS-B	1.5237	0.7216
MS-C	1.5083	0.4179
MS-L	1.6748	0.7051

In the following section, this study employs the lognormal distribution to assist in building Monte Carlo simulations of cost variance within program portfolios.

D. SIMULATION

After collating the lognormal parameters for each factor in Excel, this study leveraged the Risk Simulator add-in to perform Monte Carlo simulations for total program cost. Beforehand, however, AMA provided obfuscated baseline estimates for three Army modernization programs currently in development. At the request of the AFC commander, TRAC concealed the program names and source of data due to sensitivity. The availability of cost data through TRAC-Monterey proves paramount considering the absence of CSDR or SAR estimates on CADE due to the early-stage nature of Army modernization programs. This study refers to each of the AMA programs based on their affiliation with the closest related commodity. Therefore, the program linked to vehicles is "PROGRAM-V," the rotary-related program is "PROGRAM-R," and the C3I program is "PROGRAM-C." Using the AMA-provided notional metrics for number of formations, units per formations, unit cost, RDT&E, and procurement schedule, this study calculated the total program cost each year from 2020 to 2042 using Equation 2 before summing their individual values to obtain a baseline estimate for the entire acquisition life cycle of each program. TRAC-Monterey adjusted all values to represent base year dollars, thereby allowing the summation of the annual estimates.

$$AnnualBaselineCost = (a \times b \times c) + d \quad (2)$$

$$\begin{aligned}
&Where,\\
&a = NumberFormations\\
&b = UnitsPerFormation\\
&c = AvgCostPerUnit\\
&d = FixedCosts
\end{aligned}$$

Since no interactions exist between factors, the research focused on two separate and distinct scenarios. The first simulation forecasts the actual program costs based on the CGFs associated with the closest related commodity types. The second scenario simulates the actual program costs based on their assumed milestone events and the related CGFs associated with each milestone. All CGFs followed the stochastic parameters outlined in the previous section.

1. Commodity-based Simulation

Table 8 illustrates the basic set-up for the first scenario. The left-shaded "CGF" values represent the stochastically modeled inputs while the right-shaded "Actual Cost" values depict the resulting product of the modeled CGF and the "Baseline Estimate."

Table 8. Risk Simulator Example Set-up for Input and Outputs

Lognormal Distribution Parameters		
Factor	**Mean**	**StDev**
Vehicle	1.4707	0.5677
Rotary	1.5401	0.6571
C3I	1.4958	0.5562
Baseline Estimate	**CGF**	**Actual Cost**
$ 4,184.67	1.3044	$ 5,458.54
$ 2,680.37	2.3542	$ 6,310.07
$ 1,462.03	1.2662	$ 1,851.26
$ 8,327.07	**TOTAL**	**$ 13,619.87**

Risk Simulator automatically highlights simulation inputs with a light green fill color and simulation output values in light yellow.

After running the simulation for 10,000 iterations, the resulting interactive graphs allow the analyst to quickly query confidence level boundaries associated with a decision maker's risk tolerance. Figure 30 portrays the frequency plot on the left and the table of statistics on the right. According to Smart's research, "regardless of the underlying distribution, as long the mean and standard deviation are finite, the standard rule for exceptional variation of the mean is beyond the 80th percentile" (2021, page 173). Smart defines "exceptional variation" distributions as those that have tails that start at the 90th percentile (PCTL). In Figure 30, the red vertical line marks the 90th percentile which appears to fall directly on the start of the tail of the distribution. Subsequently, this research supports a recommendation for funding beyond the 80th percentile.

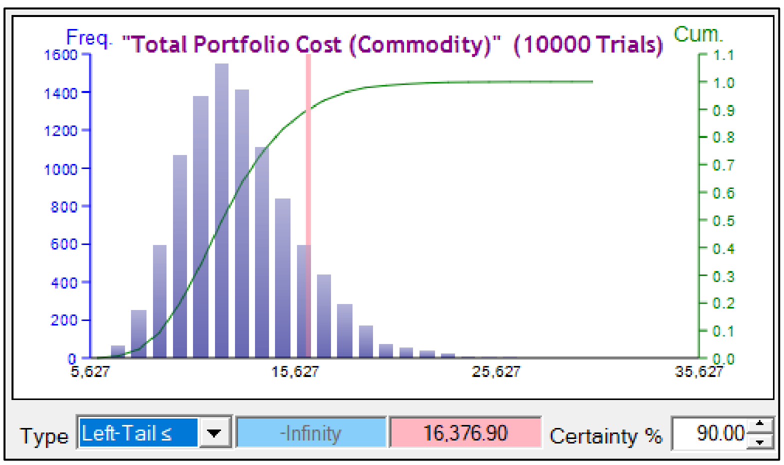


Figure 30. Risk Simulator Output for Commodity Scenario

Assuming an 80% confidence level for staying on schedule and without sacrifice of technical performance, the commodity scenario dictates a cash reserve of approximately \$9.27B beyond the initial point estimate for the entire portfolio's life cycle. The cash reserve quantity represents the monetary difference between the forecasted cost found using a left-tail confidence interval and the baseline estimate. Figure 31 highlights the stark contrast between the estimated and forecasted costs by commodity via bar chart, thereby emphasizing the reality of a cost overrun and the necessity of contingency funding.

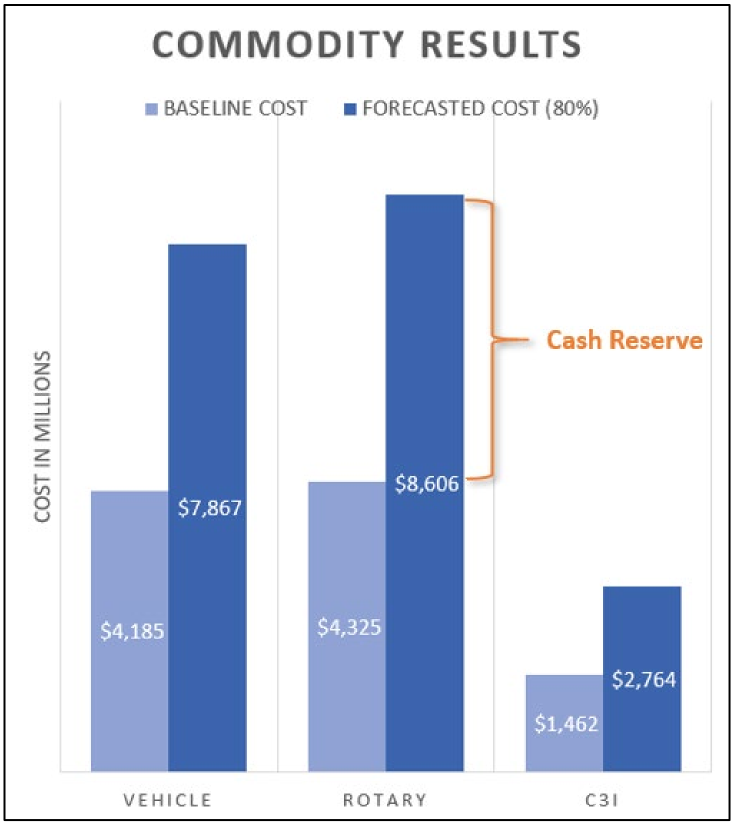


Figure 31. Bar Chart of Baseline versus Forecasted Estimates by Commodity

As a reminder, the simulation inputs are based on 102 real-world historical case studies as opposed to subjective assessments used by most risk management methodologies. For each program, the forecasted cost at 80% confidence nearly doubled after applying the historical cost variance growth factors using the Monte Carlo method. As the confidence level increases, the amount of cash reserves required to stay on-schedule and on-performance increases dramatically.

To compare the risk level associated with each program, the cost estimating industry uses the coefficients of variations (CV) (DOD and NASA 2014). Table 9 expresses the mean, standard deviation, and CV for each program as simulated by commodity. By comparison, PROGRAM-R (rotary commodity) carries the greatest risk with the highest CV at 42.9%, followed by PROGRAM-C (C3I) at 37.7%, and then PROGRAM-V at 36.7% in this simulation. When contrasted with cost magnitude and operational effectiveness scores, the CV metric can provide DMs with important prioritization insights when making programmatic decisions.

Table 9. Mean, SD, and CV Statistics for Commodity Simulation

Program	Mean (M$)	SD (M$)	CV
Vehicle	$6,149	$2,320	37.7%
Rotary	$6,666	$2,861	42.9%
C3I	$2,170	$797	36.7%

2. Milestone-based Simulation

To apply the milestone concept, the method for calculating the annual costs remained the same; however, instead of summing the entire life cycle before multiplying by the CGF, the analyst must divide the program into the phases of the acquisition life cycle as seen in Figure 32. The following assumptions for bounding the acquisition phases relate with the decision points in Figure 32.

- Milestone A begins immediately and terminates two years before initial procurement.
- Milestone B begins immediately following MS-A and terminates after the first year of procurement.
- Milestone C begins immediately following MS-B and terminates once the procurement schedule reaches 33% of the total order.
- Milestone L begins immediately following MS-C and continues in perpetuity.

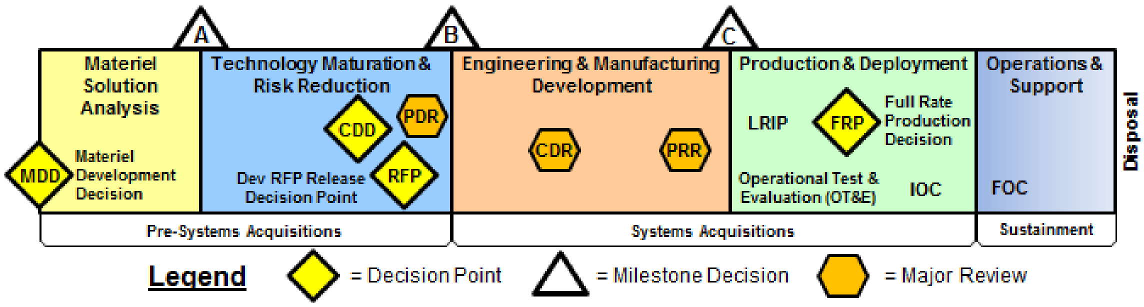


Figure 32. Acquisition Phases of a System Life cycle Diagram. Source: AcqNotes (2021).

After completing 10,000 iterations of the simulation, Figure 33 reveals the forecasted results of the total portfolio cost compared to the baseline estimate at 80% confidence since this scenario also satisfies Smart's exceptional variation criteria. That being so, the recommended cash reserve for the entire portfolio is \$9.58B which coincides with the contingency funds recommended in the commodity scenario. As before, the cash reserve quantity represents the monetary difference between the forecasted cost found using a left-tail confidence interval and the baseline estimate.

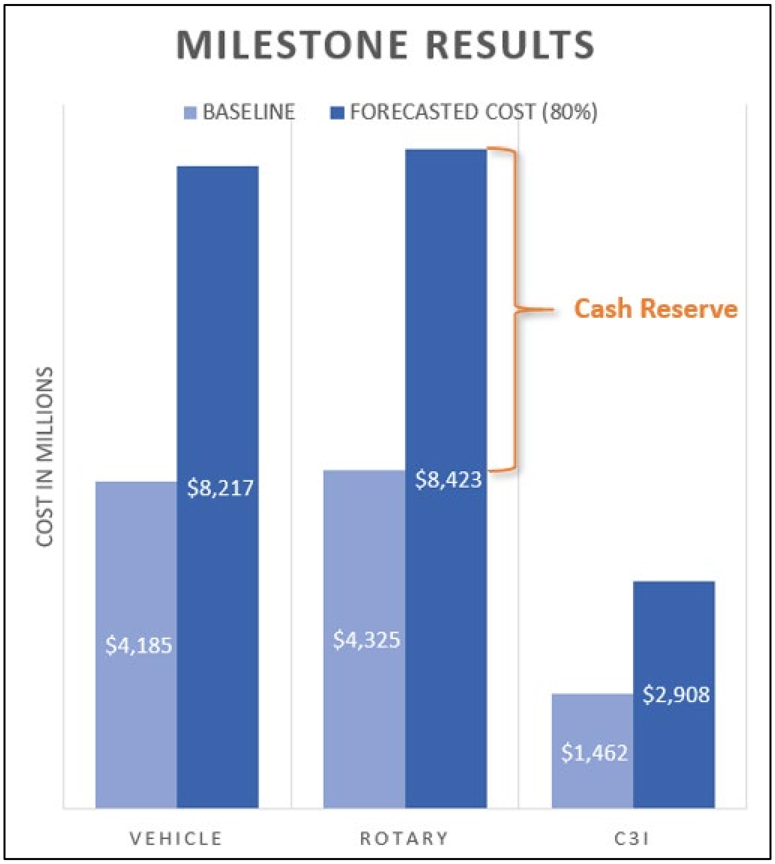


Figure 33. Bar Chart of Baseline Versus Forecasted Estimates by Milestone

The disparity between simulation outputs likely resides within the milestone assumptions that impact the start and end of each phase in the acquisition process. Since the programs in question are all pre-milestone A, the model relies on subjectivity to determine the estimated boundaries. Although the milestone assumptions attempt to follow the statutory and regulatory requirements dictated by DOD Instruction 5000.2, the actual start and end dates of programs vary dramatically. For example, although there are more than five hundred F-35s in service, the Joint Strike Fighter program has yet to reach milestone C status (Gertler 2020). The milestone-based scenario could certainly benefit from further research on event-oriented forecasting.

With respect to CV, the milestone simulation generated inversed results when compared to the commodity scenario. Table 10 displays the mean, SD, and CV for each program when modeled against their predicted milestones. In this scenario, the rotary program ranked lowest in risk level with a CV of 27.5% while the C3I program carried the highest risk with a CV of 29.9%. However, considering the minimal range between the three values, their ranking could just as easily change under the same conditions of this simulation. Overall, the milestone-based scenario forecasted near identical results at the 80% confidence level when compared to the commodity-based simulation but generated indistinguishable CV values between programs.

Table 10. Mean, SD, and CV Statistics for Milestone Simulation

Program	Mean (M$)	SD (M$)	CV
Vehicle	$6,769	$1,915	28.3%
Rotary	$6,972	$1,916	27.5%
C3I	$2,376	$710	29.9%

Considering the similarity of results between the commodity and milestone methods, this study recommends employment of the commodity-based approach over the milestone method since it maintains complete objectiveness throughout the process. Contrarily, the milestone method invites biases and inaccuracy via the necessary assumptions for determining the milestone events of a developmental program. As such, this study continues with the validation of the commodity-based methodology.

E. MODEL VALIDATION

The similarity of modeled results and exactness of deterministic outcomes between the two methods lends to the verification of both models; however, to test the validity of the simulation, this thesis relies on a five-fold cross-validation technique that compares the simulated results against a subsection of the real-world historical data from CADE. Each of the five validation simulations represent a portfolio-based scenario similar to the AMA situation at hand. An improvement in the relative accuracy between forecasted and actual data provides a mechanism for validating this study's new methodology for estimating portfolio costs.

As opposed to the hold-out method of cross-validation where a subset of the original data remains untouched throughout the entire process, the k-fold cross-validation technique maximizes statistical significance by allowing analysis on the entire dataset. Figure 34 illustrates how this study employs the five-fold method. Each row of dots numbered as "folds" represents the entire dataset used for analysis. First, this approach splits the data into test sets (20%) and training sets (80%) to achieve sufficient subset sizes and shape each scenario to reflect a real-world situation. Then, the process includes building a model against the training data before running a 10,000-replication simulation against each test set to capture the relative error as a means of determining the measure of fit. This process repeats for each fold of the validation exercise before comparing the results that ultimately determine the model's prediction performance (Seni and Elder 2010) and validity when contrasted with current methods of cost estimation.

Figure 34. K-fold Cross-Validation Technique Diagram

Based on AMA's diverse portfolio of Army modernization programs, this study assigned programs randomly to the test and training sets. This facilitates the comparison of results and highlights the operational perspective needed to validate the new methodology. Each portfolio (test set) of the five mutually exclusive scenarios included approximately twenty independent programs of mixed commodity types. After performing the Monte Carlo simulations, the results provided a computational mean for the minimum, average, and maximum percent error (PE) of the entire portfolio's cost compared to the actual observed values reported in the SAR. Equation 3 articulates the process for calculating the relative error inherent to each of the five portfolios.

$$PercentError(PE)_{forecasted} = \left| \frac{\sum_{i=1}^{n} C_{actual} - \sum_{i=1}^{n} C_{forecast}}{\sum_{i=1}^{n} C_{actual}} \right| \times 100 \qquad (3)$$

$$\begin{aligned} &Where, \\ &n = NumberofPrograms \\ &C_{actual} = TotalPortfolioActualCost \\ &C_{forecast} = TotatPortfolioForecastedCost \end{aligned}$$

Before analyzing the results of the simulations, it is important to note the relative error in AMA's existing practice. Presently, analysts use the baseline cost estimate to budget accordingly without considering the reality of life cycle cost growth when programs experience unpredictable circumstances like changing requirements or optimistic evaluations. Therefore, the relative percent error for current methodology is the total actual portfolio cost minus the total portfolio baseline estimate divided by the total baseline (merely switch forecasted cost with baseline cost in Equation 3). After running the five scenarios through 10,000 iterations, Table 11 portrays the comparison of PE results juxtaposed with the baseline method and accuracy improvement. To fully understand the range of error, the table begins (from left to right) with the baseline PE followed by the simulated minimum (MIN), 25th percentile, mean, 75th percentile, and maximum (MAX) PE before concluding with the overall percentage of improvement (benefit) when compared with the baseline using the mean PE. The bottom row of Table 11 articulates the average across all five scenarios.

Table 11. Comparison of Validation Scenario Results

Scenario	Baseline	Forecasted Percent Error					Benefit
		MIN	25th PCTL	Mean	75th PCTL	MAX	
A	34.9%	0.0%	5.7%	16.2%	22.7%	120.7%	53.6%
B	53.7%	0.0%	3.7%	9.2%	13.2%	59.0%	82.8%
C	19.5%	0.1%	16.5%	30.5%	41.4%	149.1%	-55.8%
D	36.5%	0.0%	6.2%	15.5%	22.4%	80.6%	57.5%
E	22.1%	0.1%	13.0%	24.9%	34.3%	87.9%	-12.5%
AVG	**33.3%**	**0.0%**	**9.0%**	**19.2%**	**26.8%**	**99.4%**	**25.2%**

As seen, three out of the five scenarios improved their relative accuracy based on the positive benefit values in the right-hand column of Table 11. Overall, the average improvement is 25.2%; However, this metric only captures the goodness-of-fit of the model, not the operational benefit.

To approach from the operational perspective, this study performed a Pass/Fail test on every program in each portfolio to identify those that exceeded their budgeted cost under varying constraints. This methodology proves relevant considering that when costs surpass their funding level, "the project will have to replan, rescope, reschedule, and sometime issue stop-work orders…interruptions in funding, schedule slips, and rescopes introduce inefficiencies in the project. As a result, the product is delivered later, at greater cost, and with less capability" (Smart 2021, 187); Hence, this test underscores the importance of budgeting for cash reserves. Specifically, this validity test compares the actual cost of every program in each test-set to identify those that would have theoretically exceeded their budget based on three different thresholds: (1) the baseline estimate that AMA currently uses to inform budgetary decisions, (2) the mean forecasted program cost based on the simulated results of this research, and (3) the 80th percentile forecasted cost that this study recommends. Table 12 reveals the number of programs in each scenario that surpassed their funded amounts.

Table 12. Number of Overbudget Programs by Scenario and Funding Level

		Funding Level		
Scenario	**Portfolio Size**	**Baseline Estimate**	**Mean Forecast**	**80%ile Forecast**
A	22	20	7	5
B	20	16	8	2
C	20	19	3	1
D	20	18	4	3
E	20	20	4	0
AVG	**20.4**	**18.6**	**5.2**	**2.2**

From left to right in the shaded columns of Table 12, the number of overbudget programs reduces dramatically. In every case, at least 80% of the programs in each portfolio surpassed their budgets under current practice compared to the at-most 25% level using the proposed 80th percentile of funding. The three funding level alternatives represent low, medium, and high consideration for cost overrun. In other words, budgeting at the baseline level provides no additional cash reserves other than from programs that experience cost savings, thereby inviting the most risk of program interruptions and frequent budgetary decisions. Alternatively, the higher funding level situations counteract the need to ask for more resources or sacrifice schedule or performance since they include a contingency cash reserve based on the modeled cost variance of the entire portfolio. However, there is certainly a trade-off between the number of programs that experience cost overruns and the amount of surplus money held in reserves. This methodology simply provides a mechanism for determining the appropriate balance based on the decision maker's risk tolerance when faced with resource, political, and operational constraints.

V. CONCLUSIONS

After applying a statistical approach to model the probability of cost overrun in defense contracting, this thesis proposes an improved methodology for informing decision makers on how to consider cost risk and determine contingency cash reserve levels for Army modernization portfolios. Across all services and commodities, defense contracts experience and average of 55% cost growth over the original point estimates of any given program. Decision makers must prepare for the reality of cost overrun. By leveraging newly established cost databases (CADE), analysts can now more efficiently and effectively create defendable cost estimates and aid in milestone decisions using modeling and simulations based on historical information. The result is greater confidence in cost estimates, less budgetary constraints and schedule delays, and increased technical performance.

A. RESULTS AND RECOMMENDATIONS

Based on the statistical analysis performed in this study, evidence supports the following insights and recommendations for implementing a credible, repeatable, and effectual cost estimating methodology.

- There is no statistical difference in mean CGF values between commodities, milestones, or their subsets. Therefore, *a single benchmark CGF provides relatively rapid and effective cost growth insight when under time constraints*.

- The variance within programs provides insight into their inherent risk while *coefficients of variation provide the metric for prioritizing risk levels between programs or portfolios*.

- Cost variance peaks in milestone B and specific commodities (e.g., rotary), therefore, *analysts can counter the assumption that increased cost overrun before low-rate initial production (LRIP) does not necessarily imply that the program's cost risk is escalating out of control. Further* *investigation into specific CSDR submissions can provide insight into the reasons for cost overrun.*

- Forecasting cost using the milestone approach requires subjective assumptions for determining event events of developmental programs (AMA). Consequently, this study recommends *the commodity-focused cost estimating methodology since it promotes conservatism and objectivity*.

- Analysis of historical CADE data reveals that the lognormal distribution is the best model for cost growth. When coupled with Monte Carlo simulation techniques, *cost-prediction simulations provide a sound mechanism for translating confidence levels into contingency cash reserve quantities*.

- When "exceptional variation" (Smart 2021) is present in the simulated results for total portfolio cost, the recommended funding is at or above the 80% confidence level, but *an appropriate confidence level for funding can rest somewhere between 50% and 90% depending on risk tolerance and resource availability. This mitigates the risk of costly budget interruptions that hinder schedule and technical performance.*

Implementation of these recommendations would allow cost analysts to provide a structured approach to inform budget decisions and program prioritization. The simplicity of a single benchmark CGF value manages the expectations of senior leaders while analysts provide the due diligence that accounts for historical trends. Leveraging CADE data also reduces subjectivity while streamlining the estimating process since SME-elicitation is unnecessary. Diversification of portfolios helps to mitigate the increased risk in particular commodities but relies on the ability to shift resources between programs. Lastly, modeling and simulation provides the means for quantifying the risk and unpredictability that is intrinsic to government contracting. Overall, cost growth is undeniable, so, cost estimators have an obligation to capture and communicate that truth to the relevant decision makers.

The methodology outlined in this thesis drives current practice closer toward the goal of accurate and precise cost prediction.

B. ADDRESSING THESIS OBJECTIVES

The overall approach in this study leveraged computational statistical methods of analysis to address the thesis objectives. The following summarizes the results of each objective.

THESIS OBJECTIVE #1: *Construct distributions for total program cost based on historical industry and technology maturation data.* This study equated milestone SAR data with technology maturation since it is based on the life cycle development of the overall system. After collecting, screening, and collating the CADE SAR data, this thesis applied Minitab's distribution ID plots and the Anderson-Darling goodness-of-fit test to identify the ideal distribution. This process supported the CSRUB's recommendation that the lognormal distribution is the best model for cost growth factors considering the high kurtosis and positive right skew in the data. Table 7 in Chapter IV summarized the mean and standard deviation lognormal parameters for each commodity and milestone.

THESIS OBJECTIVE #2: *Leverage Monte Carlo simulations to capture variance and confidence levels associated with cost by industry.* This study linked industry trends with commodity types to complement the context tags in the SAR data. In doing so, the variance within each commodity influenced the modeling distributions that powered the Monte Carlo simulation for total program cost. The resulting product yielded a frequency plot based on 10,000 replications that can support estimates for contingency cash requirements associated with the decision maker's confidence level. This study used the Microsoft Excel-based platform named Risk Simulator to generate forecast plots. The adjustable Risk Simulator output graphs enabled left-tailed hypothesis tests used to determine the appropriate contingency cash reserve level.

THESIS OBJECTIVE #3: *Offer data-driven information during program selection.* By calculating coefficients of variation (CV) based on the mean and standard deviation of each program or commodity's simulated results, this analysis generated comparable risk factors (CV values) that can inform program or portfolio prioritization. As an example, Tables 9 and 10 in Chapter IV articulate the resulting CV values for each commodity and milestone.

THESIS OBJECTIVE #4: *Support development of cost positions in the context of operational effectiveness*. After establishing an appropriate or assumed funding level, analysts employing this methodology can compare the forecasted cost for each funding profile with the assessed operational effectiveness (OE) benefit (OE data is available with TRAC approval but is not included in this study). A cost/OE comparison functions as a cost/benefit analysis based on the total program cost acting as the predictor to facilitate an informed decision on which alternative is best. The appropriate graph would plot the forecasted cost along the x-axis while the OE score would fall on the y-axis. Figure 35 illustrates an example of the cost/benefit graph used to identify an efficiency frontier along the periphery of the data where the alternatives generate the maximum amount of OE with the lowest possible cost. Portfolios that fall within the dominated region are exempt from consideration since they require more cost for less value. In the hypothetical case represented in Figure 35, portfolios two, four, and six are feasible alternatives since they dominate the remainder. Overall, this cost/benefit approach is a quick means for narrowing the solution space before presenting the risk and reward of the feasible alternatives.

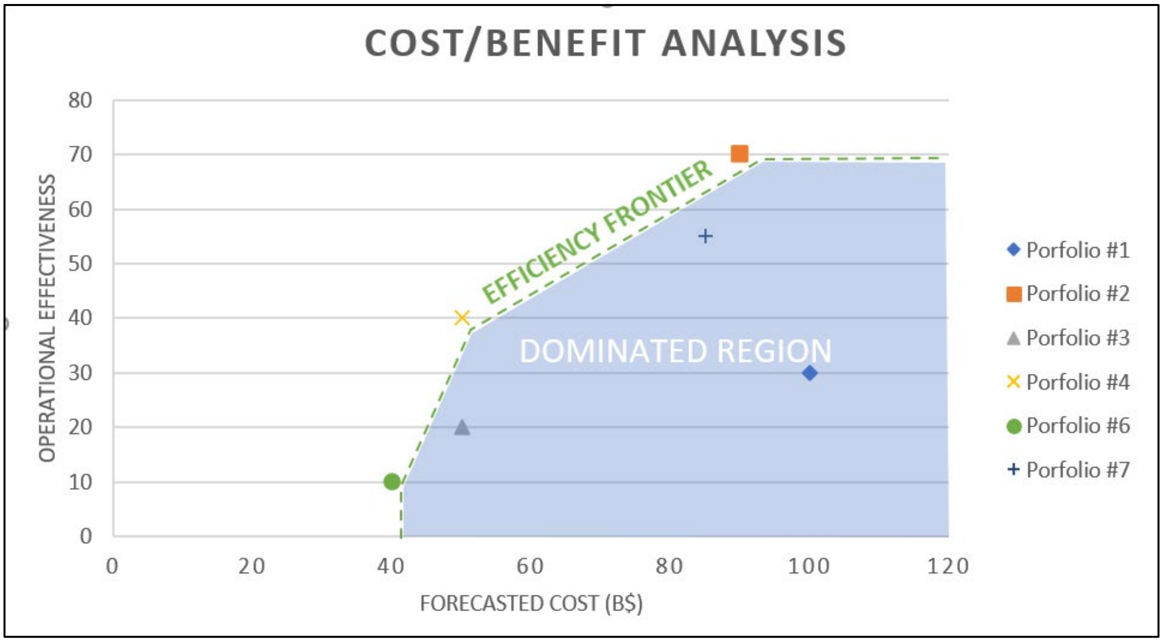


Figure 35. Hypothetical Cost/Benefit Graph

C. FUTURE RESEARCH

This study discovered that there was no statistical significance between the mean CGFs for commodities and milestones. In addition, this thesis serves to educate decision makers and analysts on the reality of cost overruns, value of historical data, and power of modeling and simulation. However, the CADE database contains more data than this study can manage. Consequently, this thesis recommends the following areas for further research.

- Develop a regression model that leverages continuous and categorical predictors such as RDT&E, engineering costs, contract-type, or outside assessments of government contractors. These factors may prove more significant in their predicting power for program specific estimates.

- Investigate the correlation between proportional costs in each of the funding categories within the baseline SAR and the final estimate. This effort will aide in developing a regression model that harnesses the predictive power of program-specific variables.

- Examine the effects of “expected shortfall” and “semi-deviation” as cost growth occurs (Smart 2021, 188). Expected shortfall offers an additional risk measure that highlights the anticipated cost and schedule growth that will occur should a program reach a specific percentile of growth (Smart 2021). Alternatively, the semi-deviation principle offers acts as another risk measure capable of identifying the cash reserve quantity should cost overruns exceed the mean (Smart 2021). Both of these risk measures would enhance this methodology by including adaptive triggers for budget adjustments.

APPENDIX A. DISTRIBUTION IDENTIFICATION PLOTS

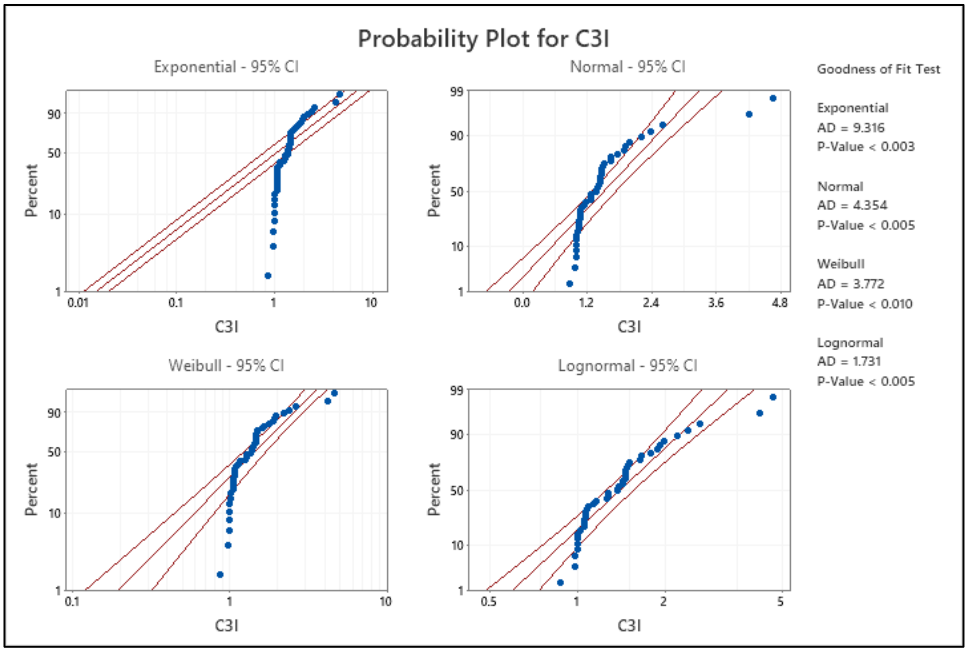


Figure 36. Minitab Probability Plot and Goodness-of-Fit for C3I Systems

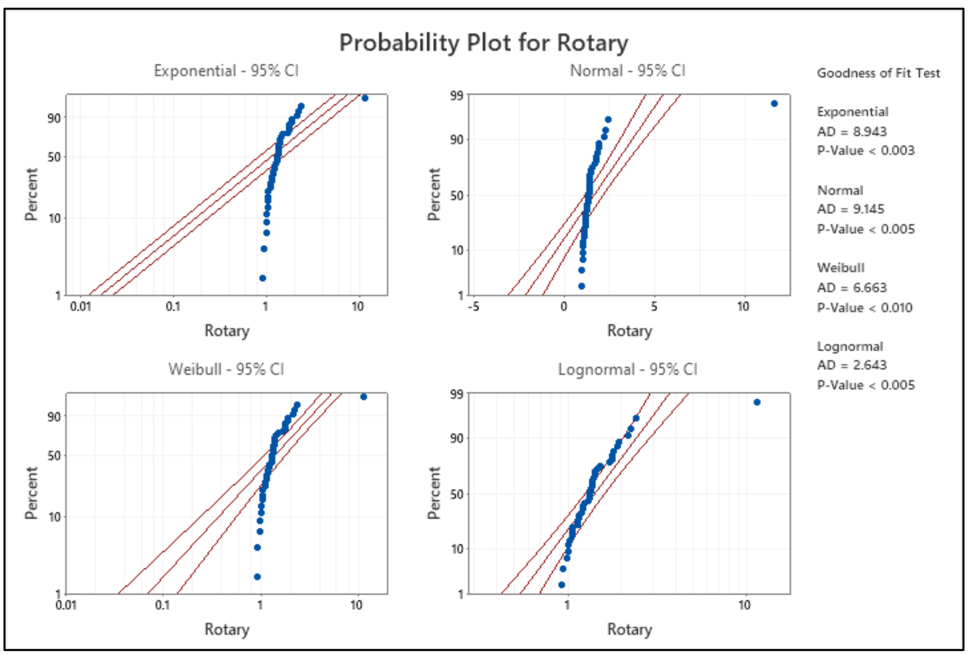


Figure 37. Minitab Probability Plot and Goodness-of-Fit for Rotary Systems

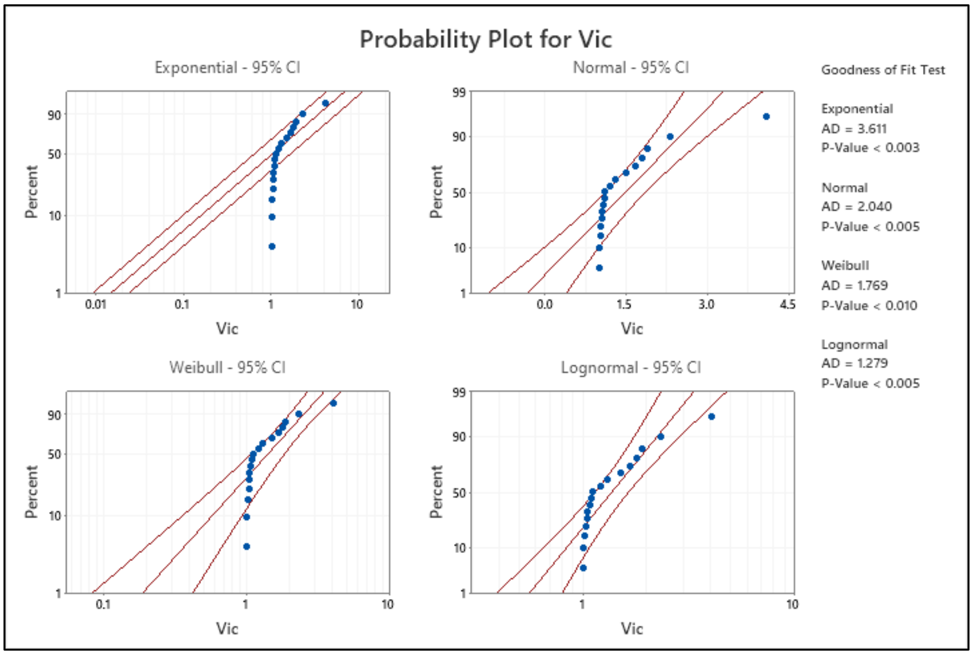


Figure 38. Minitab Probability Plot and Goodness-of-Fit for Vehicles

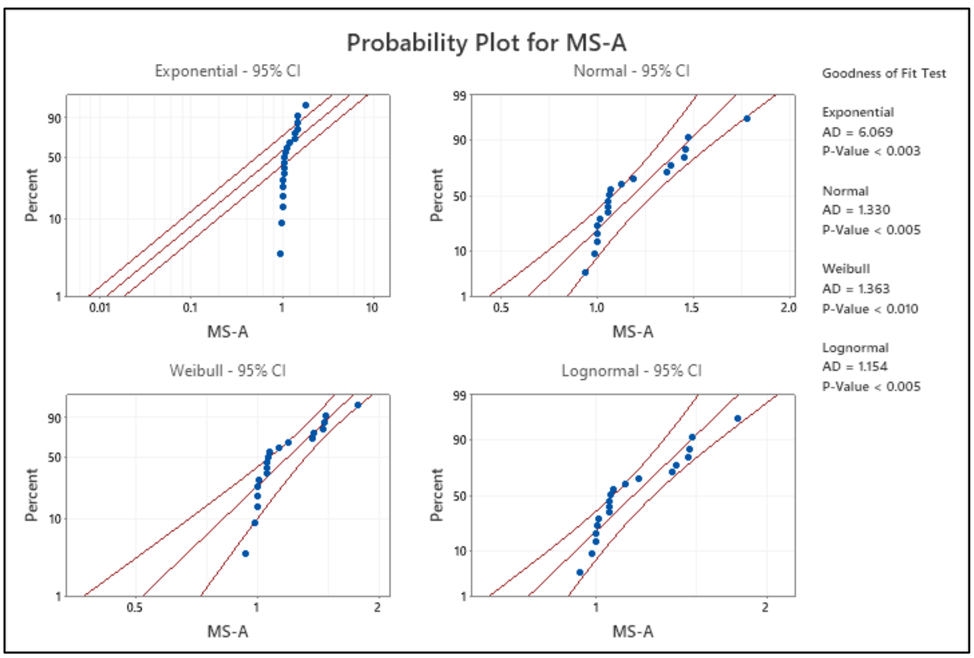


Figure 39. Minitab Probability Plot and Goodness-of-Fit for Milestone A

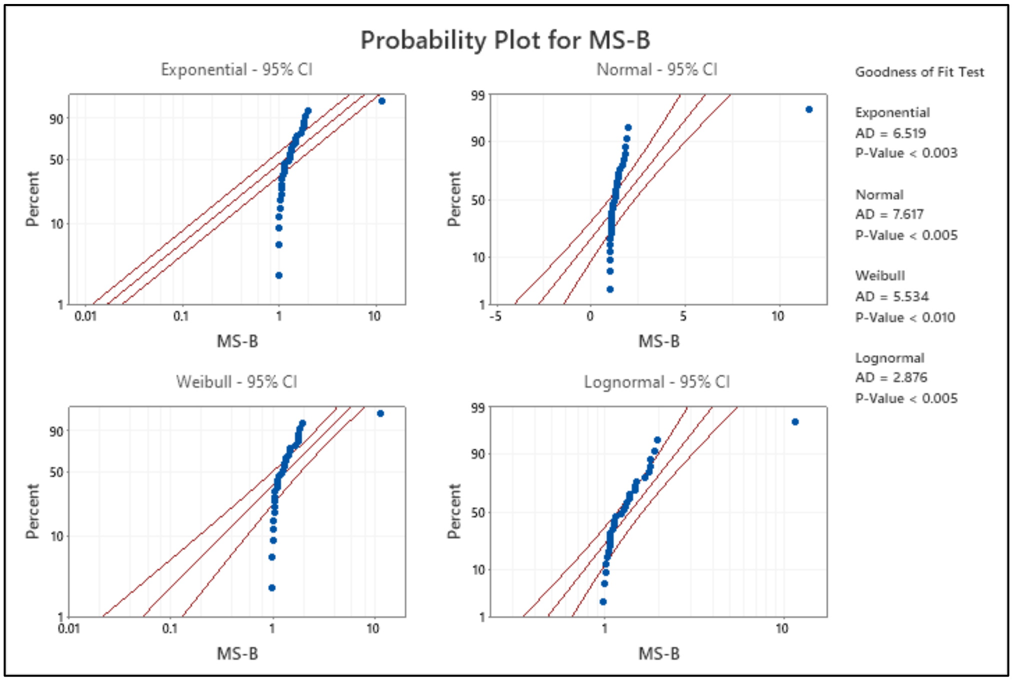


Figure 40. Minitab Probability Plot and Goodness-of-Fit for Milestone B

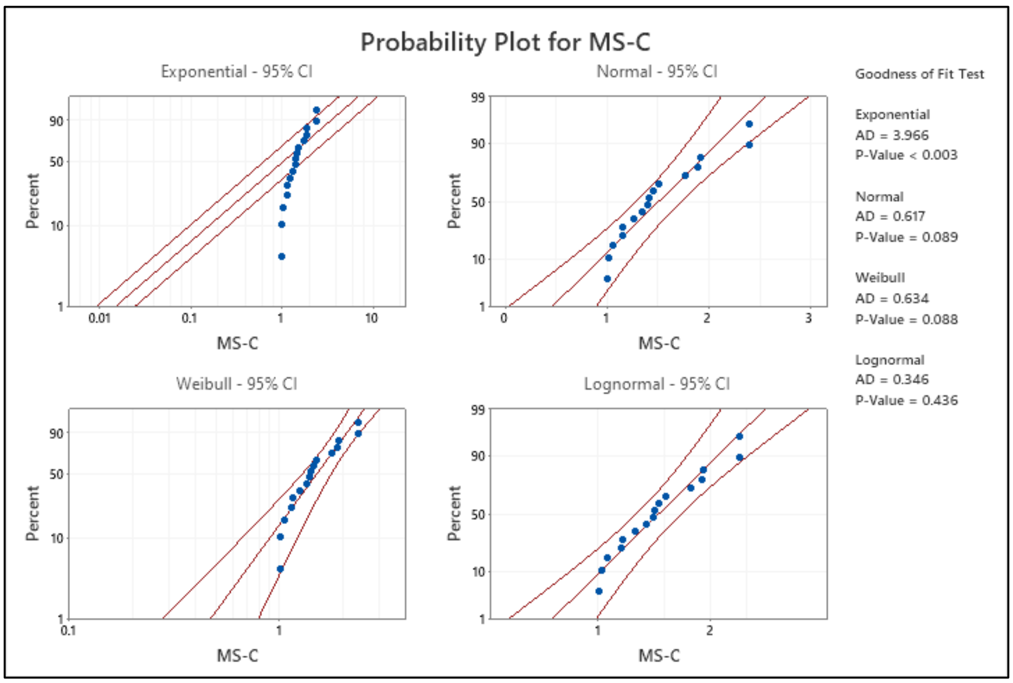


Figure 41. Minitab Probability Plot and Goodness-of-Fit for Milestone C

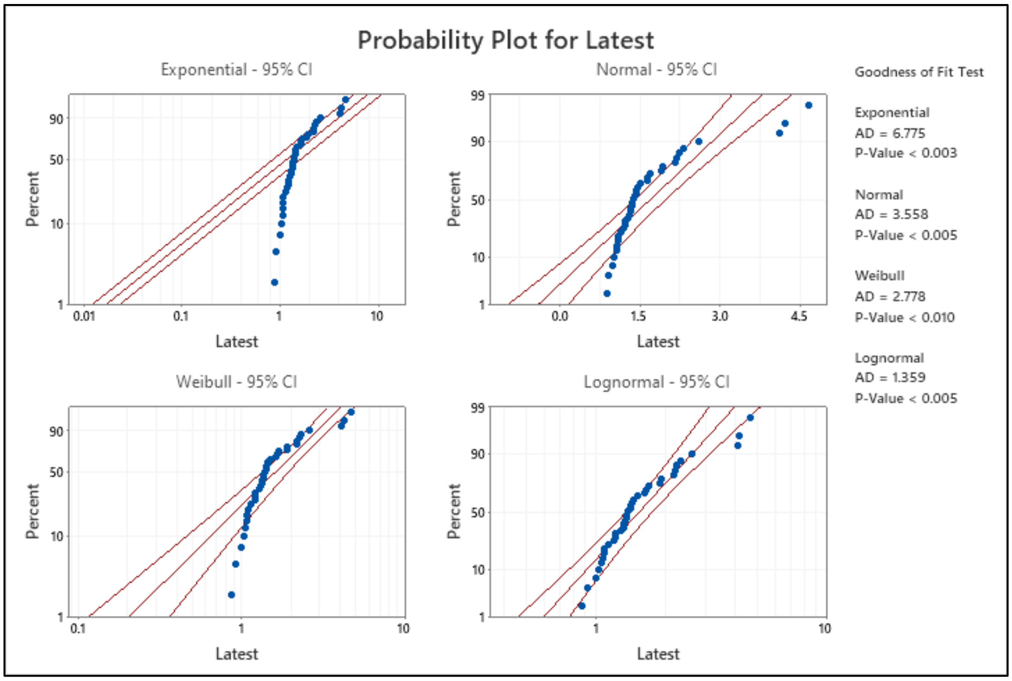


Figure 42. Minitab Probability Plot and Goodness-of-Fit for Latest SAR

APPENDIX B. DISTRIBUTION OVERVIEW PLOTS AND PARAMETERS

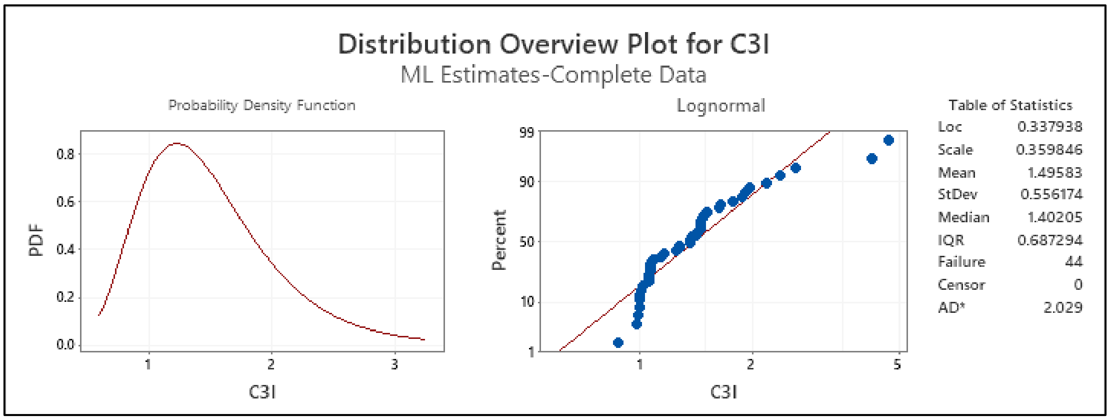


Figure 43. Minitab Distribution Overview Plot and Parameters for C3I Systems

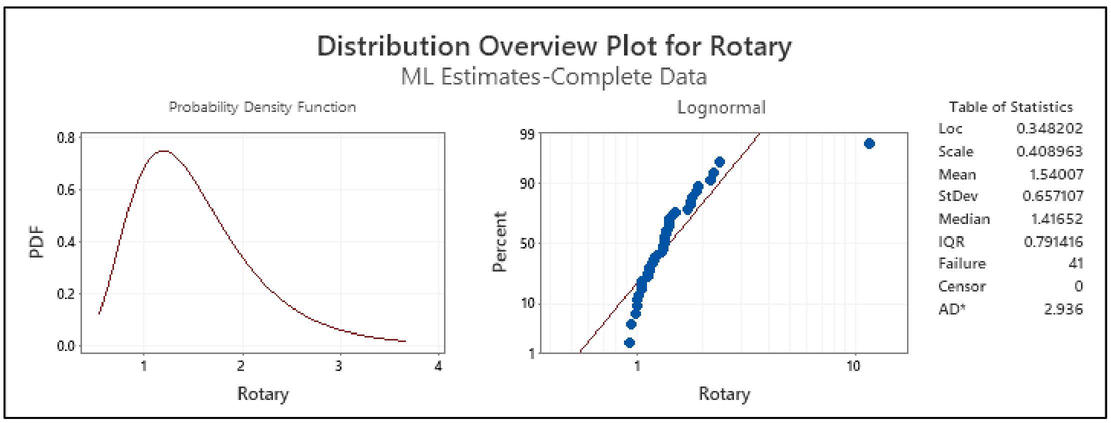


Figure 44. Minitab Distribution Overview Plot and Parameters for Rotary Systems

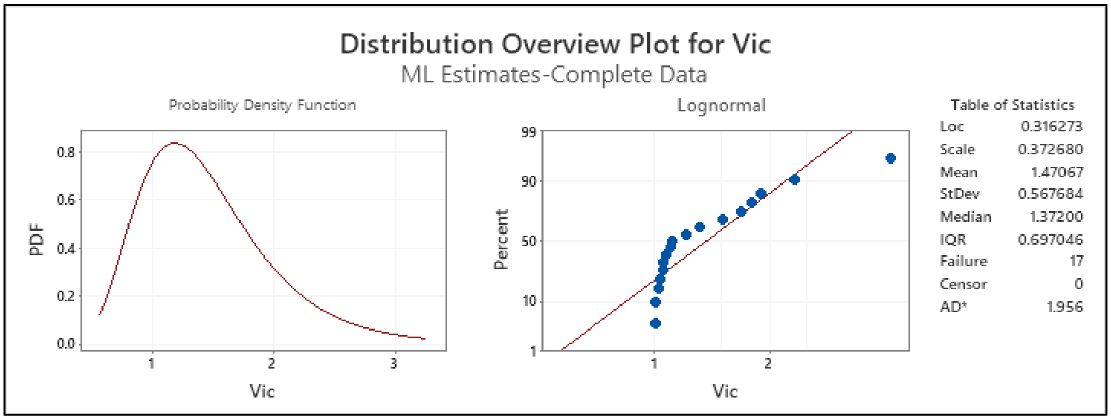


Figure 45. Minitab Distribution Overview Plot and Parameters for Vehicles

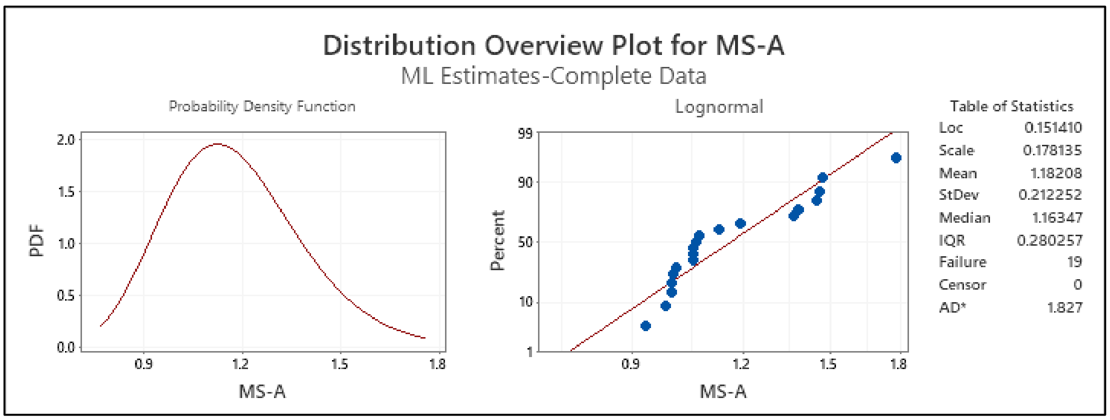


Figure 46. Minitab Distribution Overview Plot and Parameters for Milestone A

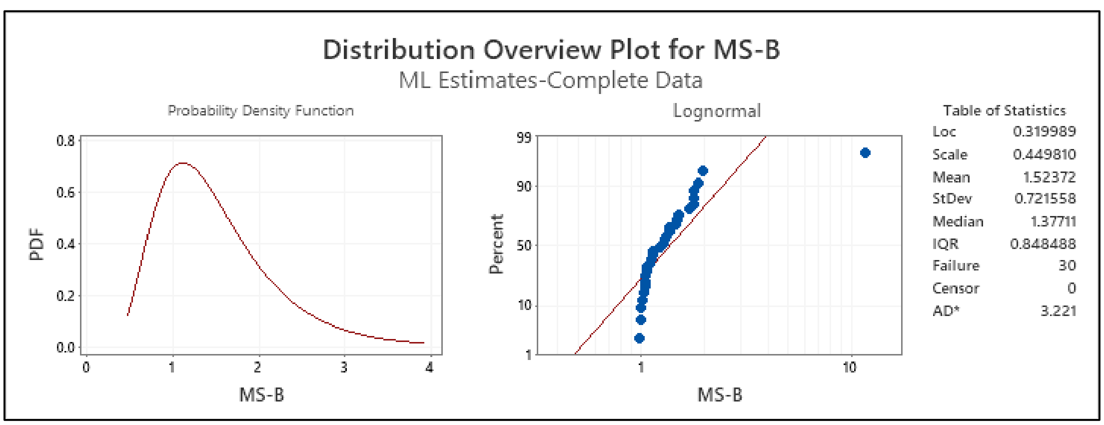


Figure 47. Minitab Distribution Overview Plot and Parameters for Milestone B

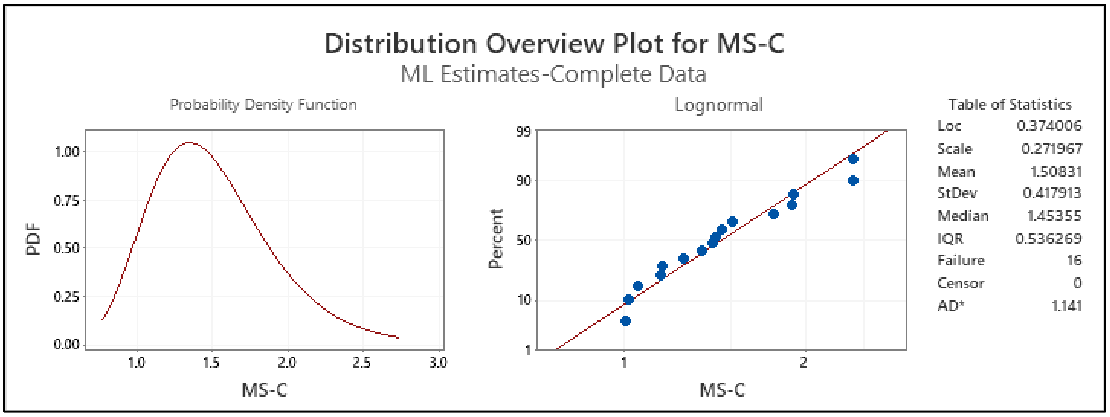


Figure 48. Minitab Distribution Overview Plot and Parameters for Milestone C

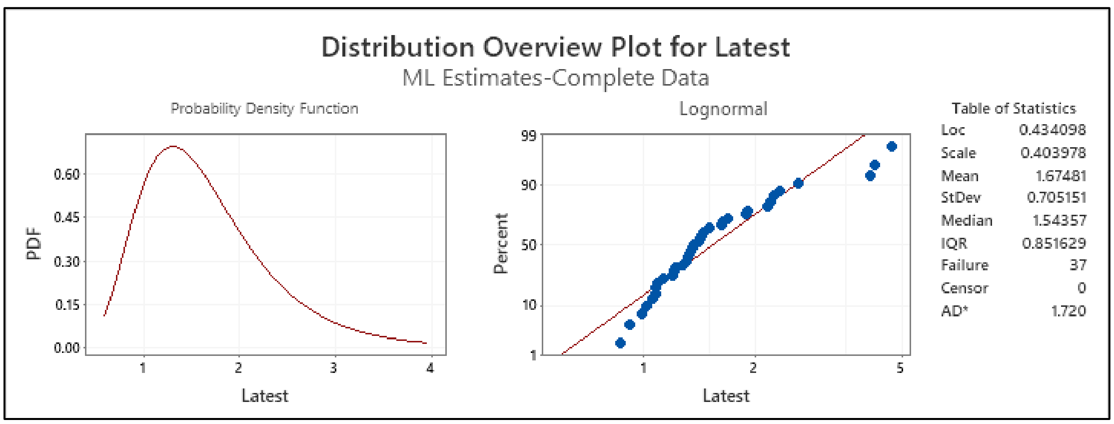


Figure 49. Minitab Distribution Overview Plot and Parameters for Latest SAR

LIST OF REFERENCES

AcqNotes. 2021. "Acquisition Phases." Last updated May 2, 2021. https://acqnotes.com/acqnote/acquisitions/acquisition-phases.

Alexander, Chuck. 2020. "Advanced Estimating Methodologies for Conceptual Stage Development." Paper presented at the *International Cost Estimating & Analysis Association Webinar Conference*, Laurel, MD: Johns Hopkins University Applied Physics Laboratory (JHU/APL). https://www.iceaaonline.com/calendar/advanced-estimating-methodologies-for-conceptual-stage- development/.

Blanchard, Benjamin S., and W. J. Fabrycky. 2011. *Systems Engineering and Analysis*. 5th ed. Upper Saddle River, NJ: Pearson Prentice Hall.

Bond, Craig A., Lauren A. Mayer, Michael McMahon, James G. Kallimani, and Ricardo Sanchez. 2015. *Developing a Methodology for Risk-Informed Trade-Space Analysis in Acquisition*. RR-701-A. Santa Monica, CA: RAND. https://www.rand.org/pubs/research_reports/RR701.html.

Department of Defense and National Aeronautics and Space Administration. 2014. *Joint Agency Cost Schedule Risk and Uncertainty Handbook*. Naval Center for Cost Analysis. Washington, DC: Department of Defense. https://www.ncca.navy.mil/tools/csruh/JA_CSRUH_16Sep2014.pdf.

Gertler, Jeremiah. 2020. *F-35 Joint Strike Fighter (JSF) Program*. CRS Report No. RL30563. Washington, DC: Congressional Research Service. https://fas.org/sgp/crs/weapons/RL30563.pdf.

Hanook, Sharoon, Muhammad Qaiser Shahbaz, Muhammad Mohsin, and B. M. Golam Kibria. 2013. "A Note On Beta Inverse-Weibull Distribution." *Communications in Statistics - Theory and Methods* 42 (2): 320–35. https://doi.org/10.1080/03610926.2011.581788.

Hulett, David T., and Bill Campbell. 2002. "3.5.5 Integrated Cost / Schedule Risk Analysis." INCOSE International Symposium 12 (1): 943–51. https://doi.org/10.1002/j.2334-5837.2002.tb02560.x.

Kansala, K. 1997. "Integrating Risk Assessment with Cost Estimation." *IEEE Software* 14 (3): 61–67. https://doi.org/10.1109/52.589236.

Kuhl, M. E., E. K. Lada, M. A. Wagner, J. S. Ivy, N. M. Steiger, and J. R. Wilson. 2009. "Introduction to Modeling and Generating Probabilistic Input Processes for Simulation." In *Proceedings of the 2009 Winter Simulation Conference* (WSC), 184–202. https://doi.org/10.1109/WSC.2009.5429329.

Lee, Richard, Peter Braxton, Kevin Concotta, Brian Flynn, and Benjamin Breaux. 2012. “SAR Data Analysis, CV Benchmarks, and the Updated NCCA S-Curve Tool.” In *SCEA/ISPA Joint Annual Conference and Training Workshop*. https://www.iceaaonline.com/2012-rsk05/.

Luher, Austin, Cody Beck, Matt Boetig, Jon Alt, Brian Wade, Andrew Ziskin, Jenifer McClary et al. 2021. “Army Modernization Analysis (AMA) Integration Team’s Development of the Trade-Space and Decision Exploration System (TRADES).” Fort Leavenworth, KS: TRAC 96.

Mislick, Gregory K., and Daniel A. Nussbaum. 2015. *Cost Estimation: Methods and Tools.* Wiley Series in Operations Research and Management Science. Hoboken, NJ: Wiley.

Mun, Johnathan. 2015. *Readings in Certified Quantitative Risk Management (CQRM): Applying Monte Carlo Risk Simulation, Strategic Real Options, Stochastic Forecasting, Portfolio Optimization, Data Analytics, Business Intelligence, and Decision Modeling.* Scotts Valley, CA: ROV Press.

Raymond, Fred. 1999. “Quantify Risk to Manage Cost and Schedule.” *Acquisition Review Quarterly*, no. Spring: 10.

Romeu, Jorge. 2003. “Anderson-Darling: A Goodness of Fit Test for Small Samples Assumptions.” *SMART: Selected Topics in Assurance Related Technologies* 10 (5): 1–6. https://www.itl.nist.gov/div898/handbook/eda/section3/eda35e.htm.

Seni, Giovanni, and John F. Elder. 2010. “Ensemble Methods in Data Mining: Improving Accuracy Through Combining Predictions.” *Synthesis Lectures on Data Mining and Knowledge Discovery* 2 (1): 1–126. https://doi.org/10.2200/S00240ED1V01Y200912DMK002.

Smart, Christian B. 2021. *Solving for Project Risk Management: Understanding the Critical Role of Uncertainty in Project Management.* New York: McGraw-Hill Education.

Wade, Brian. 2020. “TRAC-Monterey - Naval Postgraduate School.” Accessed on January 26, 2020. https://nps.edu/web/trac-monterey.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California