ERDC/ITL TR-22-1

Information Technology Laboratory

US Army Corps of Engineers®
Engineer Research and Development Center

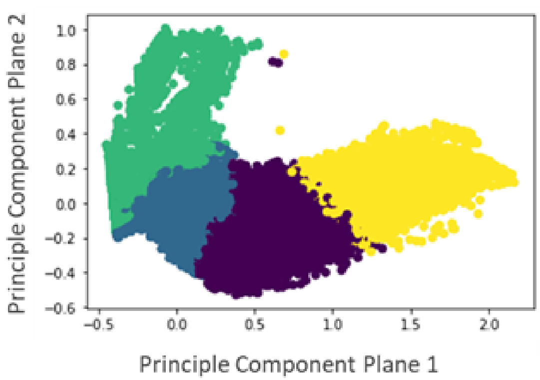
ERDC ENGINEER RESEARCH & DEVELOPMENT CENTER

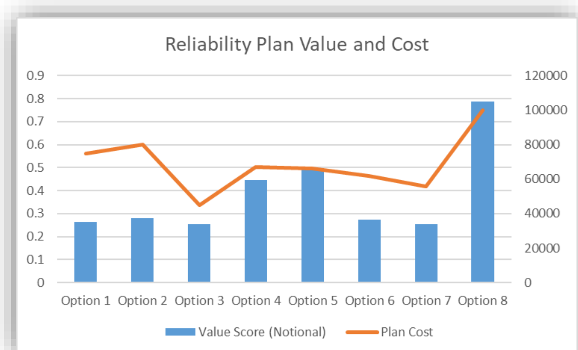
PEGSI Program: Acquisition Theory

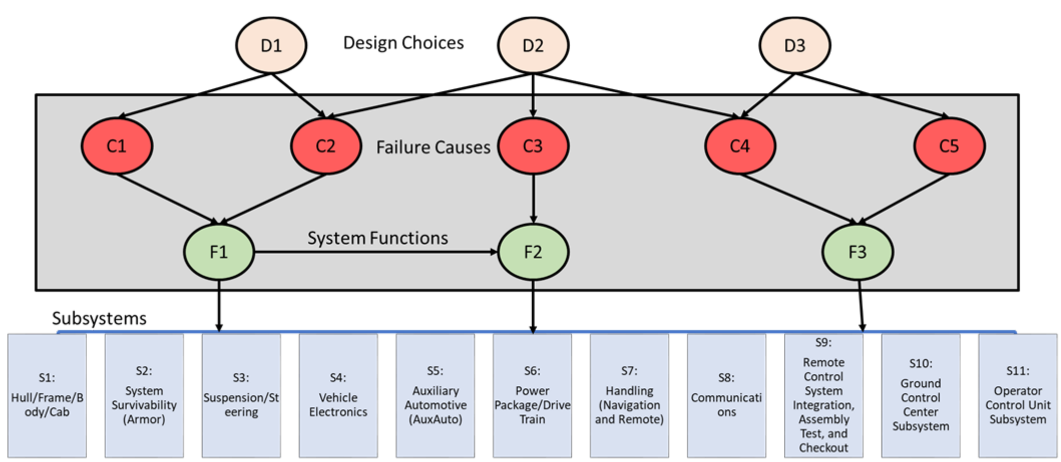
Early Life-Cycle Prediction of Reliability

Randy K. Buchanan, Christina H. Rinaudo, George E. Gallarno, and Matieu L. Lagarde

December 2022







Distribution Statement A. Approved for public release: distribution is unlimited.

The US Army Engineer Research and Development Center (ERDC) solves the nation's toughest engineering and environmental challenges. ERDC develops innovative solutions in civil and military engineering, geospatial sciences, water resources, and environmental sciences for the Army, the Department of Defense, civilian agencies, and our nation's public good. Find out more at www.erdc.usace.army.mil.

To search for other technical reports published by ERDC, visit the ERDC online library at https://erdclibrary.on.worldcat.org/discovery.

Early Life-Cycle Prediction of Reliability

Randy K. Buchanan, Christina H. Rinaudo, George E. Gallarno, and Matieu L. Lagarde

Information Technology Laboratory
US Army Engineer Research and Development Center
3909 Halls Ferry Road
Vicksburg, MS 39180-6199

Final Report

Distribution Statement A. Approved for public release: distribution is unlimited..

Prepared for Headquarters, US Army Corps of Engineers
Washington, DC 20314-1000

Under Army Direct Funding, Program Element Number 0603463A, Project Number: BP4, "Acquisition Theory—Reliability."

Abstract

The intent of this project is to investigate a variety of approaches for the development of a basic model for the early life-cycle prediction of reliability (pre-Milestone A). The United States Department of Defense (DoD) currently utilizes an acquisition framework in which system development advances through a series of checkpoints known as milestones. Each milestone represents a decision point, with Milestone A being the earliest in the life cycle. At Milestone A, also known as the risk-reduction decision, the DoD evaluates design concepts while also committing funds to the maturation of technologies in an effort to mitigate future risks. Typically, little is known about the particular system to be developed at this point in the acquisition life cycle, but DoD regulations require program managers to submit system reliability information (OUSD[A&S] 2015). Traditional reliability predictions, however, require extensive knowledge of the system of interest to produce accurate results. This level of knowledge is unavailable at or before Milestone A; therefore, there is a need to create models and methodologies for the prediction of system reliability. This report provides an overview of a variety of methods investigated to improve the prediction of early life-cycle reliability.

DISCLAIMER: The contents of this report are not to be used for advertising, publication, or promotional purposes. Citation of trade names does not constitute an official endorsement or approval of the use of such commercial products. All product names and trademarks cited are the property of their respective owners. The findings of this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

DESTROY THIS REPORT WHEN NO LONGER NEEDED. DO NOT RETURN IT TO THE ORIGINATOR.

Contents

Figures and Tables

Figures

Tables

Preface

This study was conducted for the US Army Engineer Research and Development Center under Program Element Number 0603463A, Project Number: BP4, "Acquisition Theory—Reliability." The technical monitor was Dr. Mihan McKenna.

The work was performed by the Institute for Systems Engineering (ISER) of the Computational Science and Engineering Division (CSED), US Army Engineer Research and Development Center, Information Technology Laboratory (ERDC-ITL). At the time of publication, Dr. Simon R. Goerger was director, Institute for Systems Engineering; Dr. Jeffrey L. Hensley was chief, Computational Science and Engineering Division; and Dr. Robert M. Wallace was the technical director. The deputy director of ERDC-ITL was Dr. Jackie S. Pettway, and the director was Dr. David A. Horner.

COL Christian Patterson was the commander of ERDC, and Dr. David W. Pittman was the director.

1 Introduction

1.1 Background

The intent of this project is to investigate an approach for the development of a basic model for the early life-cycle prediction of reliability (pre-Milestone A). The United States Department of Defense (DoD) currently utilizes an acquisition framework in which system development advances through a series of checkpoints known as milestones. Each milestone represents a decision point, with Milestone A being the earliest in the life cycle. At Milestone A, also known as the risk-reduction decision, the DoD evaluates design concepts while also committing funds to the maturation of technologies in an effort to mitigate future risks. Typically, little is known about the particular system to be developed at this point in the acquisition life cycle, but DoD regulations require program managers to submit system reliability information (OUSD[A&S] 2015). Traditional reliability predictions, however, require extensive knowledge of the system of interest to produce accurate results. This level of knowledge is unavailable at or before Milestone A; therefore, there is a need to create models and methodologies for the prediction of system reliability.

1.2 Objectives

Objectives include the development and demonstration of a reliability modeling approach for early reliability prediction as well as implementation of an initial early reliability functionality for trade-space analysis . The reliability and cost modeling will be integrated while incorporating a set-based design (SBD) framework.

Why is this important? Reliability is critical because it contributes to a system's warfighting effectiveness as well as its suitability in terms of logistics burden and the cost to fix failures. Furthermore, DoD regulations require project managers (PM) to submit system reliability info before Milestone A because it impacts performance, cost, and schedule estimates.

Why it is difficult? Field operational data on the same or similar products are the best estimate of a product's reliability but are difficult and expensive to collect. New designs have no directly relevant data. Traditional reliability predictions require extensive knowledge of the system, generally unknown early on. Per DoD 5000.02, the PM must formulate a comprehensive reliability and maintainability (R&M) program using appropriate strategy to ensure R&M requirements are achieved (OUSD[A&S] 2015).If the methodology includes cost, schedule, and performance impacts, it can assist a program manager with tradeoff decisions.

1.3 Approach

This research approach includes developing a model through quantifying reliability at an appropriate level of abstraction (subsystem) and aggregating into system-level estimates while acquiring historical defense systems data. Process/model results are compared to historical defense systems' data. Models also illustrate the impact of reliability values upon the system's costs throughout the life cycle (specifically spares and maintenance—two of the major drivers) given the predicted reliability values are modeled.

2 Reliability Prediction Overview

Initial research activities included a review of DoD established instructions and guidebooks that provided summaries of key activities and requirements associated with evaluating reliability during the materiel acquisition process. This referencing of established DoD guidelines is important to developing new concepts that may have to be integrated into the existing infrastructure.

2.1 DoD reliability requirements

The Reliability Key System Attribute (KSA), which is a portion of the mandatory Sustainment Key Performance Parameter (KPP), is defined as follows: “Reliability measures the probability that the system will perform without failure over a specified interval under specified conditions. Reliability must be sufficient to support the warfighting capability needed in its expected operating environment. Considerations of reliability must support both availability metrics. Reliability may be expressed initially as a desired failure-free interval that can be converted to a failure frequency for use as a requirement” (Office of the Secretary of Defense in collaboration with the Joint Staff 2009).

The Reliability KSA along with the Maintainability and the Operating and Support Cost help to “ “ensure that the Sustainment KPP is achievable and affordable in its operational environment. Together, the KPP and supporting KSAs ensure early sustainment planning, enabling the requirements and acquisition communities to provide a capability solution with optimal availability, maintainability, and reliability to the warfighter at an affordable life cycle cost” (Defense Acquisition University 2022).

Additionally, the *Department of Defense Reliability, Availability, Maintainability, and Cost Rational Report Manual* provides a methodology to use when considering reliability, availability, maintainability, and cost (RAM-C) requirements, aids in the development of realistic requirements, and ensures that sustainment is considered early in the program life cycle (Office of the Secretary of Defense in collaboration with the Joint Staff 2009).

2.2 Reliability metrics review

Review of reliability metrics included discrete and continuous systems. Discrete systems included single-use items like munition, where reliability is commonly measured as probability. Methods include probability density function, cumulative distribution function, reliability function, and failure rate function. Continuous use systems include aircraft, vehicles, etc. Reliability for these systems is commonly measured in terms of a primary usage metric (operating hours, miles, flight hours, etc.). Example methods include Mean Time Between Failures (MTBF), Mean Time to Repair (MTTR), Mean Time to Maintenance (MTTM), Maintenance Down Time (MDT), and exponential distribution.

Many methods exist to predict reliability of existing systems; however, the prediction of reliability during pre-Milestone A acquisition activities introduces challenges due to the conceptual nature of the system where many of the design components are unknown. A high percentage of DoD systems struggle to meet reliability requirements (Brady 2013). Providing a model with a capability to more accurately estimate early conceptual system reliability could allow the use of reliability as a metric earlier in the decision-making process. Adoption of this model could result in the development, manufacture, and fielding of systems with high reliability. Increased system reliability results in decreased system failure, thus increasing system availability. If a system does not achieve good reliability, mission performance and operations and sustainment costs are at risk. End-to-end modeling of the system life cycle helps to evaluate the impact of changes in reliability.

The fidelity of reliability analysis increases for a system as it progresses throughout the acquisition life cycle. Reliability estimations in the pre-Milestone A phase rely primarily on comparisons to legacy weapon systems. The actual reliability performance of a fielded system is often used to inform earlier phases of reliability analysis and could also be used in modeling and simulation to inform design decisions. The use of modeling and simulation is more prevalent in the early phases of system development. Once prototypes are developed, the prototype data are used for reliability analysis. As the system matures to full-rate production, the information gathered from the field is primarily used for reliability analysis. Throughout the system's life cycle, reliability and sustainment impacts total ownership costs.

2.3 Current DoD reliability estimation methods

From the *DoD Guide for Achieving Reliability, Availability, and Maintainability*, existing reliability assessment methods require the use of system, subsystem, and component information to construct relevant mathematical models capable of assessing reliability, availability, and maintainability (DoD 2005). This component information comes from a variety of sources, including subject matter experts and system documentation. Analysts use the collected information to inform inputs for chosen reliability estimation techniques.

Common reliability estimation techniques include the following:

- Reliability block diagram modeling techniques
- Monte Carlo simulations
- Markov analysis (Markov Chain and Markov Process)
- Failure modes and effect analysis (FMEA)
- Fault tree analysis (FTA)
- Ishikawa diagram

The assumptions used by system engineers and other subject matter experts directly affect the accuracy of the aforementioned techniques. To help guide future efforts in determining relevant avenues for pre-Milestone A reliability estimation, and the subsequent development of in situ correction factors, a more detailed analysis of modeling assumptions is necessary.

2.3.1 *Reliability block diagrams*

Reliability block diagrams are at the core of initial pre-Milestone A reliability predictions. Reliability block diagrams describe the interaction of system components within a given system. These system block diagrams allow for description of both components that operate in series and in parallel. Each block, then, calculates the reliability of a system component. The reliability of the system is then calculated using the reliability of all the system components and component interactions. Because of the robust nature inherent to reliability block diagrams, this technique continues to be used to estimate overall system reliability. For pre-Milestone A reliability estimates, however, block reliability estimates must be calculated using physics-based methods or using statistical information for similar systems. Statistical estimates can be further improved if the system under development uses previously deployed system components. The results of these models can then be integrated into system-level simulation. The potential for incrementally improving a system block diagram reliability estimate makes this technique highly recommended as a starting point for improving in situ (and post hoc) reliability estimates, though its suitability for large-scale physical modeling is questionable.

2.3.2 *Physics-based reliability (PBR) modeling*

Development of novel systems inherently limits the availability of system data and necessitates the utilization of physics-based modeling techniques. Finite element analysis (FEA) is a commonly used method for mechanical systems, though alternative physics of failure techniques and mathematical models are employed for this purpose for different system types. Techniques such as FEA allow for component engineers to assess reliability of their component when exposed to real-world stress. Aggregating the component-level, physics-based reliability estimates can allow for system-level reliability estimates. However, prior findings illustrate common pitfalls for these predictive techniques for reliability estimation (Littlefield, Mazzuchi, and Sarkani 2012; Nguyen and Zheng 2013; Fiondella and Zeephongsekul 2011; Thodi, Khan, and Haddara 2013; Fiondella and Xing 2014) as well as provide possible paths to incrementally improve existing techniques (Hu and Du 2018), including the development of in situ correction factors for specific types of systems (e.g., ground vehicles or rotorcraft).

2.3.3 *Monte Carlo simulation and Markov analysis*

The development of physics-based models at the component level allows for an estimation of the likelihood of component failures. The statistical information pertaining to each component can then be integrated into a Monte Carlo simulation environment or used for Markov analysis. While these techniques are commonly used in reliability assessments, it is worth noting that the output of these techniques is only as good as the inputs that are provided. Thus, if physics-based modeling techniques provide poor estimates for likelihood of failure, the poor estimates compound upon each other to negatively impact the estimation of system-level reliability produced through Monte Carlo simulation or Markov analysis. As such, reassessing, verifying, and validating the assumptions used for physics-based models provides an avenue for improving current DoD early lifecycle reliability prediction techniques.

2.3.4 *Failure modes and effect analysis and fault tree analysis*

While Monte Carlo simulation and Markov analysis can identify potential sources of system failure, FMEA and FTA are more commonly used to identify potential sources of failure. These two techniques are typically used together to ensure a more thorough assessment of system failures. Specifically, reliability engineers typically complement the forward inductive logic of FMEAs for a bottom-up system assessment with the deductive logic of the FTA for top-down assessment in order to account for all major sources of system failures. The accuracy of the FMEA and FTA is contingent upon the system-level information collected from component and system engineers. Verification and validation of the system assumptions could result in better identification of failure sources, translating into an improvement of reliability estimation capabilities.

2.3.5 *Ishikawa diagrams*

Ishikawa diagrams show the cause and effect of some hypothetical event for a specified system. While great for helping organize thoughts on system reliability at a high level, the Ishikawa diagram has a tendency to oversimplify causes of system failure. While potentially useful at a component or subsystem level, Ishikawa diagrams are poor tools for analyzing complex systems. The insufficiency stems from the Ishikawa diagrams' inability to characterize relationships between causes of failure. In order to help improve reliability estimation, reassessment of the conditions used for Ishikawa diagrams could correct inaccuracy in reliability estimates.

The methods the DoD leverages to estimate reliability are

- Traditional system engineering series and parallel/redundant reliability modeling techniques (DoD 2005, 4-15–4-16),
- Monte Carlo simulations (DoD 2005, 4-17–4-18),
- Markov analysis (Markov chain and Markov process) (DoD 2005, 4-18–4-19),
- Failure modes and effect analysis (DoD 2005, 4-23–4-25),
- Fault tree analysis (DoD 2005, 4-25–4-27),
- Ishikawa diagram (DoD 2005, 4-27–4-28), and
- Benchmarking (DoD 2005, 4-28–4-29).

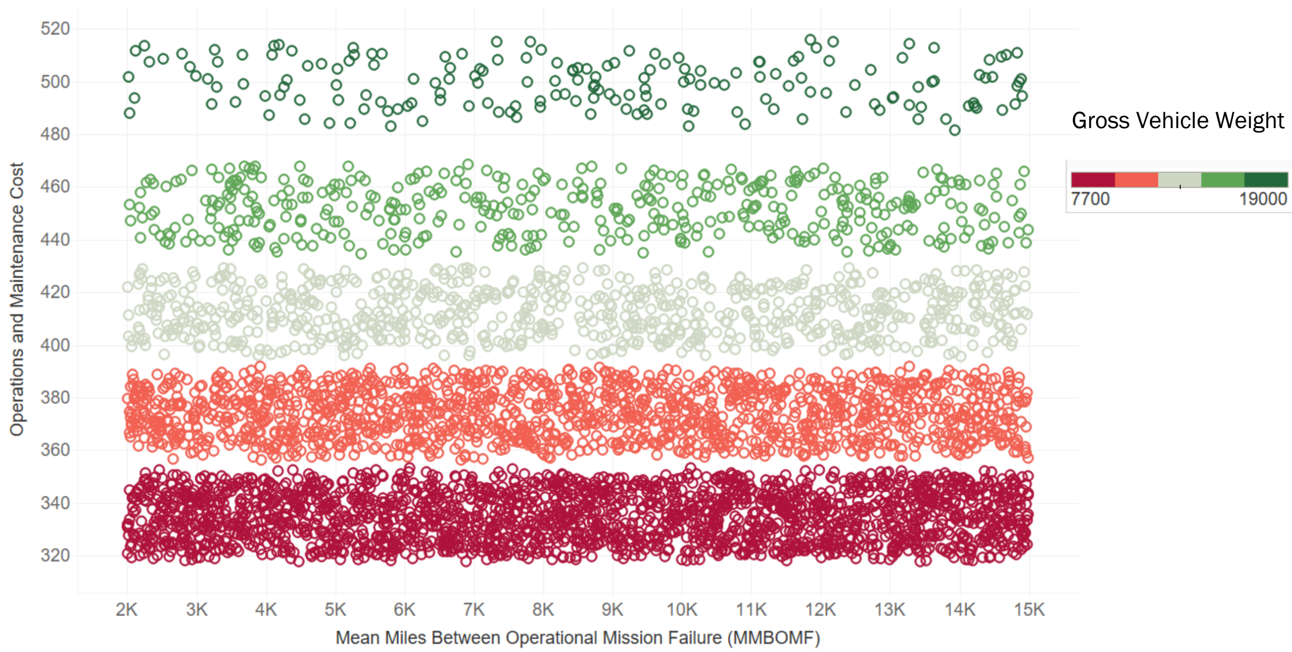
These techniques assume that conditions such as failure rates of components follow a homogenous exponential process. They are based off observational data (covariate, component, system, and subject-matter-expert judgement), inductive logic (demonstrably incomplete), or the existence of similar systems. Additionally, these techniques can require extensive systems knowledge and long computational runtimes (such as during Monte Carlo simulations), thereby delaying reliability assessment efforts. Correcting the assumptions of existing techniques would provide a possible path to estimating the true reliability of a system (or subsystem).

It should be noted that the above techniques are used throughout the development process, with early life cycle techniques being contingent on the type of system under development. If previous systems (or subsystems) are utilized to construct a new system, previously developed empirical, statistical, and physics-based models can be utilized to establish correction factors for the evolutionary (sustaining) system.

2.4 Linking reliability to life-cycle cost (LCC) pre-Milestone A

Reliability factors impact a program's life-cycle cost (LCC). This research effort investigated the impact of reliability upon pre-Milestone A LCC estimates. An investigation of the reliability variable in a case study of the Light Reconnaissance Vehicle cost model revealed that the operations and maintenance (O&M) cost component within the cost model was not sensitive to variances in the Mean Miles Between Operational Mission Failure (MMBOMF). Instead, the cost model was sensitive to the vehicle weights. The model output (Figure 1) depicts how the model's O&M cost generates O&M cost bands related to the vehicle weights (note that the colors corresponding to notional vehicle design weights of 7,700; 10,387; 13,073; 15,760; and 19,000) and therefore are not sensitive to the changes in the MMBOMF. The banding in the figure results from the five chassis designs considered for analysis. Future research could investigate how MMBOMF impacts the associated maintenance, sustainment, and logistics cost estimates.

Figure 1. MMBOMF versus estimated operations and maintenance cost.

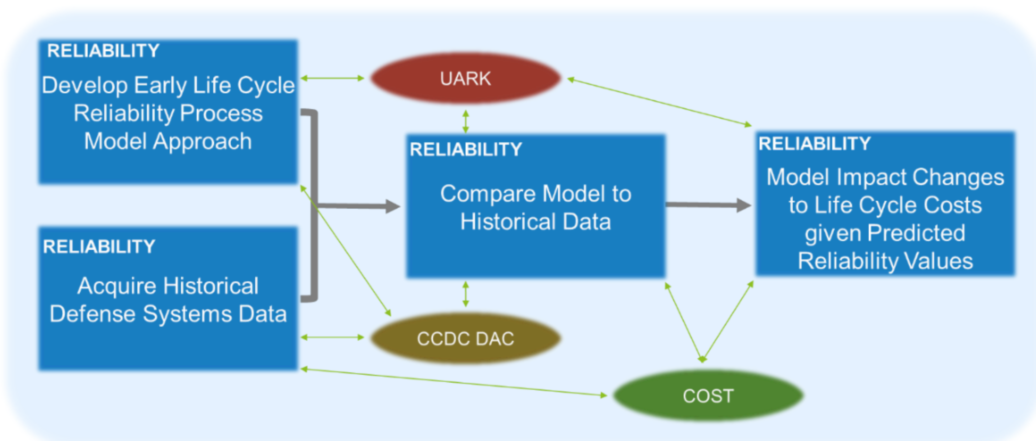


3 Technical Approach Overview

3.1 Methodology overview and workflow

The early life-cycle prediction of reliability research approach includes developing a model through quantifying reliability at an appropriate level of abstraction (subsystem) and aggregating the information into system-level estimates. The research leverages historical defense systems data. The approach will model impact changes to LCC (specifically spares and maintenance—two of the major drivers) given the predicted reliability values. A diagram of the approach and collaborators' roles is shown in Figure 2.

Figure 2. Reliability task workflow.



Collaborators included the Combat Capabilities Development Command (CCDC) Data and Analysis Center (DAC) and the University of Arkansas (UARK). UARK assisted with a detailed literature review, investigations into Bayesian hierarchical models, parametric models, reviews of reliability growth modeling techniques (used for life-cycle analysis), and definitions for a modeling approach linking models to an SBD framework. CCDC DAC assisted with developing a process for early life-cycle predictions at pre-Milestone A, comparing model results to historical defense systems' data, and modeling impact changes to life-cycle costs given the predicted reliability values. The research integrates cost analysis at appropriate steps during the entire development process.

3.2 Methodologies

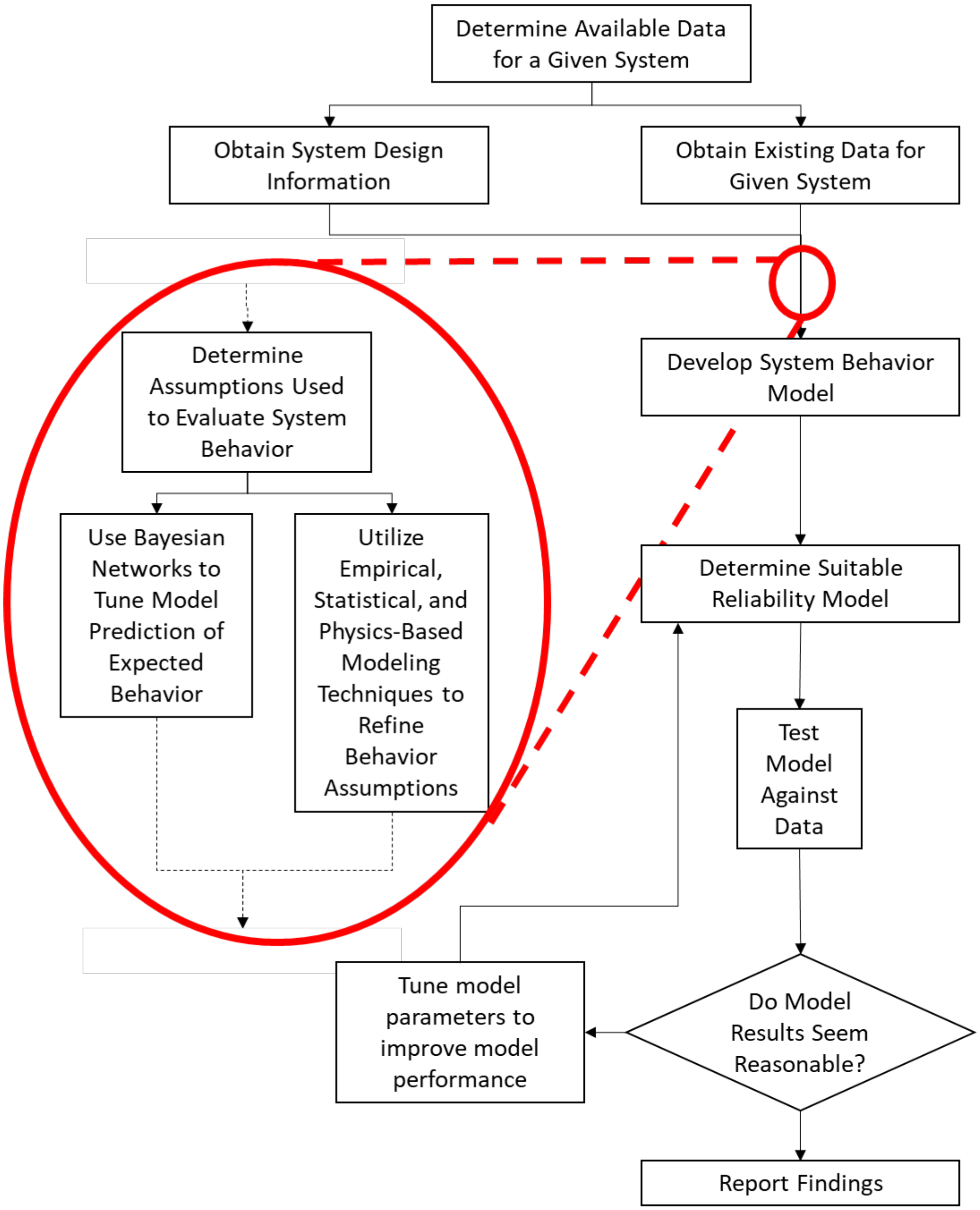
Current DoD RAM assessment techniques were reviewed. In summary, RAM assessment approaches leverage three categories of mathematical models: empirical, statistical, and physics-based (DoD 2005). The existing methodologies have significant gaps in predictive abilities for early life-cycle analysis. The research team considered two possible approaches to mitigating this issue: (1) integrate in situ correction techniques or factors to improve descriptive, predictive, and prescriptive analytic capabilities; and (2) apply post hoc corrective factors to target the gap between predicted and empirical reliability values.

Principal approaches to model development include in situ, post hoc, and program metadata crosswalk analysis. All of these approaches are dependent on data sources and are discussed in the following sections.

3.2.1 In situ approach

The integration of in situ correction factors for existing reliability models' parameters allows for potential improvement in descriptive, predictive, and prescriptive analytic capabilities (Figure 3). This approach is advantageous since it provides corrected results upon model output; however, correcting existing models requires extensive time and effort to verify and validate the modification for each mode. Further drawbacks to this method include the need for extensive system knowledge when looking to correct and improve physics-based models, large data sets, and detailed analysis when correcting empirical and statistical models. As such, implementation of techniques such as Bayesian networks could ease the burden for analysts and provide expedient updates to some existing models. Overall, the in situ approach is relatively difficult to implement but is a robust modeling technique.

Figure 3. In situ methodology.



3.2.2 Post hoc approach

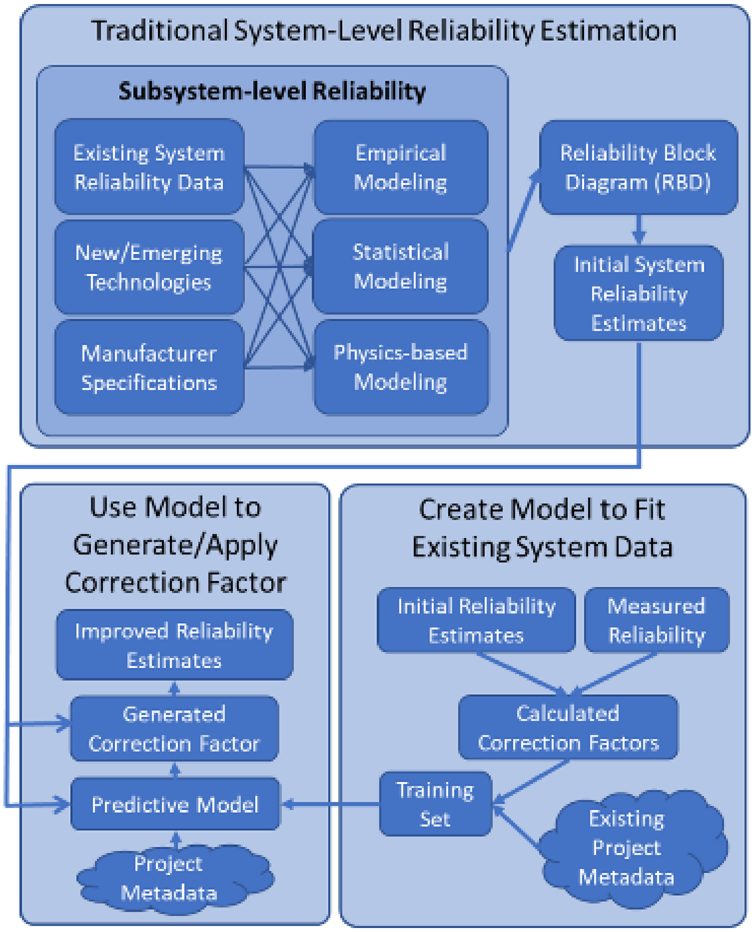
One method of addressing the gap between predicted and empirical reliability values in the existing methodology would be the application of post hoc corrective factors. This technique should generate an improved overall reliability analysis in the early stages of a project (pre-Milestone A). This method applies post hoc corrective factors to target the gap between predicted and empirical reliability values (Figure 4). The difference between predicted and measured reliability for existing systems would be used to generate a set of corrective factors for reliability estimation.

A potential methodology for accomplishing this would be to create a post-processing system to improve upon the baseline metrics provided by traditional reliability predictions. By taking the baseline metrics and combining them with metadata related to the overall project and project management information, it should be possible to produce a per-project correction factor that can be used to adjust the initial estimate to better match empirical reliability and improve overall predictive accuracy.

In order to produce such a model, it would first be necessary to obtain full life cycle documentation for several DoD acquisition projects and analyze the predicted reliability provided within the reports at various stages of the process, from pre-Milestone A to end-of-life. Once this has been done, analysis can be performed per-project to determine the direct relationship between predicted reliability and measured reliability in the form of a correction factor by incorporating variables that are not otherwise taken into account by traditional reliability prediction analysis. Which variables are included and the associated weights of those variables would be adjusted in order to best fit the existing data for inputs and correction factors determined through analysis of existing projects. After the model has been produced and is well fit to the existing data, additional project information can be used as a testing set to verify the effectiveness of the predictive abilities of the model.

The primary advantage of this method is that it is relatively simple to implement when compared to some of the alternatives and is easy to iterate upon since it is relatively isolated from existing systems. Another advantage is that it is easy to add to an existing process as it does not require modifications to the underlying process and, instead, seeks to correct errors after the fact. However, because it does not directly identify and solve underlying problems within the methodologies it builds upon, it has a more limited ability for correction.

Figure 4. Post hoc methodology.



3.2.3 Program metadata crosswalk approach

An effort was undertaken to acquire and analyze scorecard data at the program level to extract and determine impacts to reliability. Considerations for model integration included CCDC DAC-provided input, tools, and data, as well as cost research and data.

This approach includes examining available CCDC DAC data from the Army Materiel Systems Analysis Activity (AMSAA) Reliability Scorecards. These Scorecards were compared against Director, Operational Test and Evaluation (DOT&E) reports, cost system data, and available ground vehicle data sets. The advantage of this research method is that it provides a holistic look at available data for systems across the DoD acquisition life cycle. However, the limitations are related to gaining access to specific system data as well as the time-intensive research to review and crosswalk system information that may or may not exist in a variety of collected data.

3.2.4 Integrated framework approach

Design decisions are integrated into the reliability model to capture the impact of those decisions on reliability. The reliability models and capabilities of those design decisions will then be used to assess performance requirements and calculate value and LCC of design alternatives. This approach leverages previous research using a ground vehicle cost model file.

4 Data Collection

4.1 Data sets available

Several sources of data were leveraged to support this research effort. Data sets used for all models are shown in Table 1.

Table 1. Model data sources.

Principle Data Sources				
Data	Platform	Data Type/Format	Size	Source
Reliability Scorecards	Various	Excel	4.66 MB	CCDC DAC
DOT&E Annual Reports	Various	PDF / Text	434 MB	DOT&E Website
Sensor Data and Maintenance Logs	Air Vehicle	HPC Database	~70 TB	ERDC ITL
Sensor Data	Ground Vehicles	HPC Database	~70 TB	ERDC ITL/CCDC DAC

4.2 AMSAA Reliability Scorecards

The CCDC DAC provided access to AMSAA Reliability Scorecards that are used to examine and identify reliability best practices during system development. The data provided included scores and ratings for systems previously evaluated by CCDC DAC. Scorecard criteria categories are as follows:

- Reliability Requirements and Planning
- Training and Development
- Reliability Analysis
- Reliability Testing
- Supply Chain Management
- Failure Tracking and Reporting
- Validation and Verification
- Reliability Improvements.

While the Scorecard criteria are not all applicable to pre-Milestone A—Failure Tracking and Reporting, for example—there is still useful information contained within the first three criteria categories. Specifically, the Reliability Requirements and Planning section helps identify which tools and methods have been used during system development. The Training and Development category gives information on whether the project team is using industry standard and newly developed techniques to assess system reliability. Lastly, the Reliability Analysis category rates a project's application of reliability analysis methodology. Information from these Scorecards will help inform prioritization of in situ correction factor development as well as identify sources of risk commonly found in early life cycle system development. (CCDC DAC, n.d.)

4.3 Director Operational Test and Evaluation annual reports

Although reliability metrics are generally difficult to gather without specific system access, DOT&E reports provide a source of project reliability data on many DoD systems. DOT&E annual reports documenting system reliability information can augment metadata-modeling efforts, thus improving fidelity. There are 20 years of DOT&E annual reports available for analysis (2000–2019). Reports identify high-level reliability metrics for numerous transportation systems (ground vehicle, aircraft, and watercraft) and weapons platforms (Office of the Director, Operational Test and Evaluation, n.d.).

4.4 Air vehicle data set

The air vehicle data set contains a variety of flight and maintenance data for various deployed rotorcraft and includes raw sensor data and maintenance logs. The main data set is quite large (>60 TB in size) and accessible through Engineer Research and Development Center's (ERDC) high performance computer (HPC) platform (with approval from the appropriate data point of contact). As a result, any potential methods of processing these data for reliability assessment will be limited to HPC-centric approaches. However, the HPC Data Analytics group was able to provide access to a small subset of simulated sensor data for use outside of the HPC environment for initial analysis.

Preliminary analysis of the sample data provided shows that, while not directly usable as a source of reliability information, the breadth and depth of the data has the potential to allow for determining system reliability at a subsystem level (and in some cases component level). This data set has already been used previously for failure analysis. The data set was investigated for potential use in this research but was not used due to the significant processing requirements.

4.5 Ground vehicle data set

Access to a US Army ground vehicle data set composed of 45 different vehicles with over 1 million operational days, collectively, was made available for analysis. Data for the ground vehicles were collected through a controller area network (CAN) bus and consist of 60 columns at 1 value per second. Previous research efforts examined semi-supervised and unsupervised machine learning methods by which to explore and analyze the data set (Bond et al. 2020). Discussions with the Bond research team included topics related to the previously researched supervised and unsupervised methods, the extent to which these methods were considered during the previous model development process, and data analysis scripts used to explore the ground vehicle data set. Similarities exist between the previous hybrid approach research (Bond et al. 2020) and the post hoc approach considered by this research project.

Initial analysis of the data set required access to the HPC system as well as some custom software scripts to load and parse a subset of the available data in order to provide statistics and indices for use in the development and selection of techniques for further analysis. The information included a list of the platforms within the data set and a list of available fields for each of those platforms. Additionally, this analysis produced the number of systems of each platform for which there are available data and the total number of operational days available for each platform. Table 2 displays a subset of the data generated through the initial analysis.

Table 2. Subset of data generated through initial analysis of the ground vehicle data set.

Parameter	Value
Total Number of Platforms	107 Platforms
Total Number of Systems	4,628 Systems
Average Systems per Platform	43 Systems/Platform
Total Operational Days	~1.06 Million Operational Days
Average Operational Days per Platform	~10,000 Operational Days/Platform
Average Operational Days per System	~230 Operational Days/System
Total Number of Parameters	12,991 Parameters
Average Number of Parameters per Platform	121 Parameters/Platform
Total Size of Data set	6.8 Terabytes (Minified JavaScript Object Notation [JSON])
Average Amount of Data per Platform	~64 Gigabytes/Platform (Minified JSON)
Average Amount of Data per System	~1.5 Gigabytes/System (Minified JSON)

This information helped with contextualizing the amount, form, and type of data, and allowed for comparing the available systems with the systems found within the DOT&E reports in order to determine candidate systems for further analysis and model development.

5 Model Implementation

Based on available data, the research implemented a hybrid model, based on in situ and post hoc methodology, a value model, and an integrated model.

5.1 Hybrid model implementation

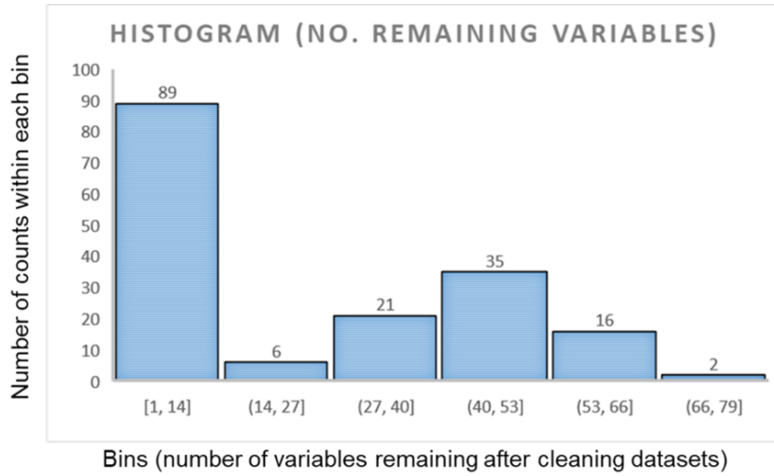
5.1.1 Ground vehicle sensor data

Before leveraging the ground vehicle data set, it was necessary to examine the quality of the data set in order to determine fitness for use. For the selected ground vehicle system there were 169 time series data sets available (Figure 5). Each data set contains data recorded at the 1 Hz (per second) unit level. Initial data cleaning efforts examined column variables to determine which variable had missing values (Not a Number or NaNs) for greater than or equal to 20% of the data set. For other variables, interpolation was used to estimate missing values. The most frequent number of variables remaining after quality assessment was three variables, one of which was the time index. Inconsistency of usable variables presented an issue for time series and regression model development. It should be noted that the number of variables remaining should not be interpreted as the same variables remaining across all data sets (Figure 6). For example, a cleaned data set with a set of 40 variables does not necessarily intersect with another cleaned data set with a smaller set of variables (e.g., 13 variables).

Figure 5. Data set summary statistics.

SUMMARY STATISTICS	
Mean	21
Median	9
Mode	3
Std Dev	21
Min	1
1st Quartile	3
2nd Quartile	9
3rd Quartile	41
Max	69
Range	68
No. VINs above 2nd Q.	84
No. VINS above 3rd Q.	20

Figure 6. Number of variables remaining after cleaning data sets



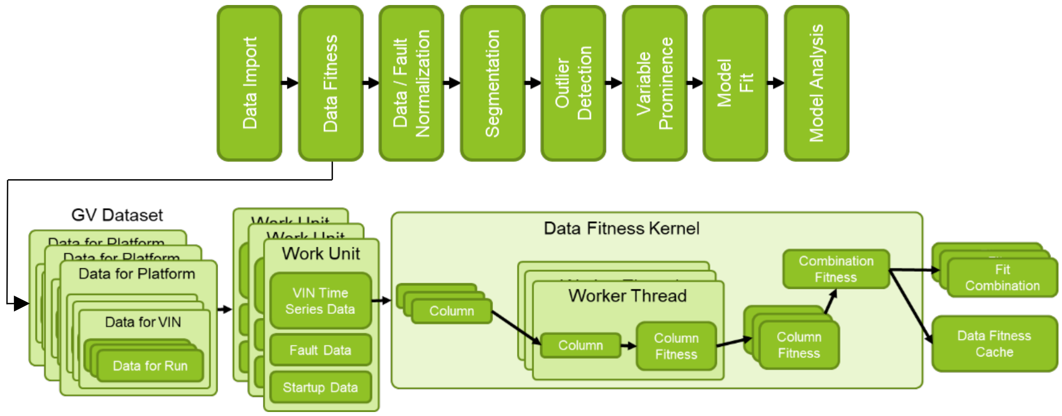
5.1.2 Data analysis

Data analysis techniques for the hybrid model implementation included dimension reduction, cluster analysis, regression analysis, and the application of decision trees. First, cluster analysis was conducted on ground vehicle data that were subject to dimension reduction by principle component analysis, t-distributed stochastic neighbor embedding (tSNE), or uniform manifold approximation and projection (UMAP). This data set was then subject to k-means and density-based spatial clustering of applications with noise (DBSCAN) cluster analysis. Next, regression analysis was performed in an attempt to estimate response variables within the data set using simple linear regression, multiple linear regression, and smoothing splines. Lastly, decision tree models were generated using the extreme gradient boosting (XG Boost) method.

5.1.3 Transition to the HPC

Efforts to implement the hybrid model included modifying model implementation to achieve speedup on the HPC (Figure 7). This included splitting model components into highly parallel subcomponents. The outcome was a speedup (12×) when compared to performing analysis on individual combinations locally. Using the HPC to run the model on the entire data set achieved an additional speedup (35×) and improved output quality. The pipelined, caching architecture allowed for faster iterative model development and avoided duplication of computational tasks between consecutive runs locally and on the HPC, which enabled more flexible use of HPC time and increased flexibility of data importing. The pipelined, caching architecture was derived after analyzing sample code from the research of Bond et al. (2020). The deployment of this architecture allowed for an easy transition of the model to other data sets (e.g., other ground vehicles, air vehicles, etc.)

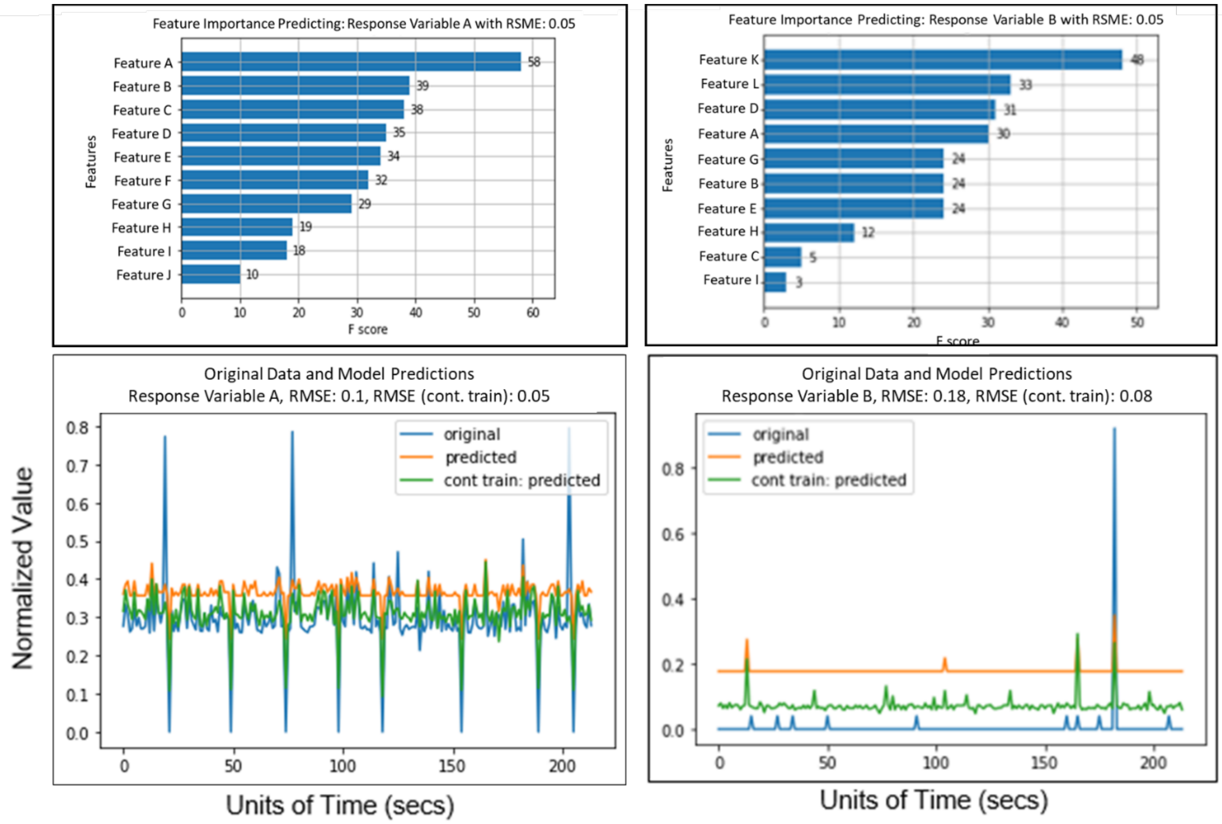
Figure 7. Hybrid model transition to HPC.



5.1.4 Hybrid model outcomes and analysis

The XG Boost regression (boosted tree method) provided the best model performance and results. The model regressed with squared loss (which more heavily penalizes errors on extreme values) and provided the most flexibility for model improvement through continued machine learning. The results of the regression provided insights into the ground vehicle under assessment, revealing correlations and impacts between variables. Key findings involved two response variables (A and B). For 71% of the time, an outlier was present before fault on the response variable B. Similarly, for 80% of the time, an outlier was present before fault on response variable A. Figure 8 describes the feature importance and time series prediction for the response variables A (*left*) and B (*right*) with XG Boost.

Figure 8. Hybrid model outcomes and results.



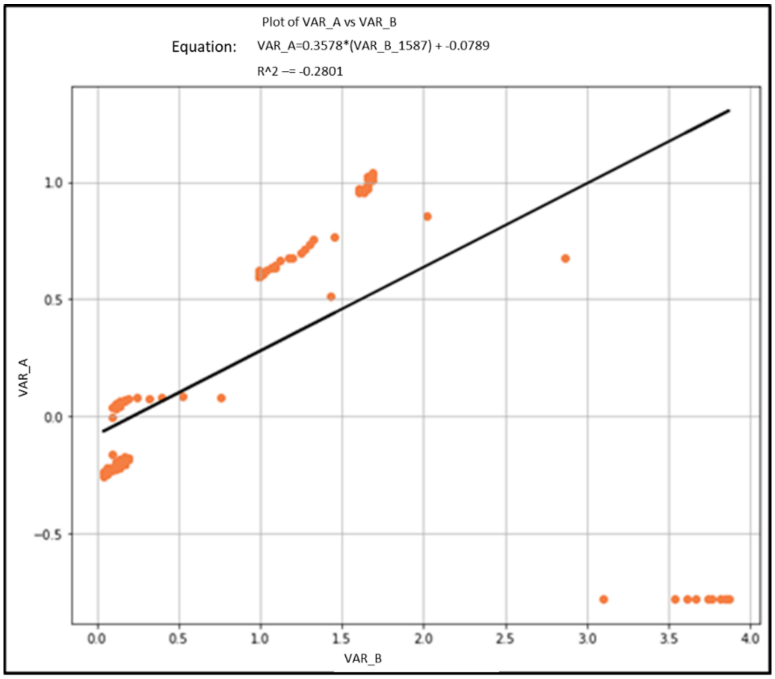
5.1.5 Data analysis script development

Development of data analysis scripts focused on incorporating previous coding efforts to automate the fitting of simple linear regression, multilinear regression, and smoothing spline models to time series segments as well as XG Boost models for entire time series.

5.1.5.1 Simple linear regression

The simple linear regression analysis utilized time series segmentation, where, for a given VIN, the time series was segmented using fault data. Each time series segment was analyzed for outliers, and if an outlier was detected, all variables with outliers during the time series segment and within 500 units after the identified fault were used to subset the time series data frame for the segment. All combinations of the noted variables were examined, and a simple linear regression model was fit. Results were saved to a data frame that can be output as a comma-separated value (CSV) file for further analysis, and a visualization was generated (Figure 9).

Figure 9. Example simple linear regression plot on a time series segment data set.



The output file includes data on the following: dependent variable, independent variable, slope, intercept, mean absolute error, mean square error, root mean square error, and *R*-squared value (Figure 10).

Figure 10. Example simple linear regression analysis of ground vehicle time series data.

Index	X-Var	Y-Var	Slope	Intercept	Mean Abs Error	Mean Sq Error	Root Mean Sq Error	R Sq
0	X Var A	Y Var A	-0.904925	1.40559	0.528492	0.339578	0.582733	0.594048
1	X Var A	Y Var B	0.309126	-0.373825	0.314837	0.116765	0.341709	0.468936
2	X Var A	Y Var C	0.089554	0.0886566	0.43297	0.225901	0.47529	0.0531253
3	X Var A	Y Var D	0.20882	-0.15584	0.109158	0.0585922	0.242058	0.535567
4	X Var A	Y Var E	1.14664e-32	-0.301358	0.0238312	0.0863251	0.293811	-0.00662252
5	X Var B	Y Var F	-0.522712	1.14128	0.338818	0.266363	0.516104	0.56498
6	X Var B	Y Var G	0.0063719	-0.139846	0.355681	0.225891	0.47528	-0.0273851
7	X Var B	Y Var H	0.30631	-0.0589546	0.336625	0.258353	0.508284	-0.0828996
8	X Var B	Y Var I	-0.0754812	0.0586608	0.15311	0.105108	0.324204	0.166856
9	X Var B	Y Var J	1.47648e-33	-0.301358	0.0238312	0.0863251	0.293811	-0.00662252
10	X Var C	Y Var K	0.77422	0.876268	0.347244	0.334528	0.578384	0.453654
11	X Var C	Y Var L	0.0276279	0.711195	0.72499	0.844589	0.919015	-0.00967232
12	X Var C	Y Var M	1.03347	0.297614	0.256597	0.0990756	0.314763	0.584719

Analysis of the output file could yield potential equations for modeling the relationship between two variables. Then, by examining a dependent variable and independent variable pair across the recorded time series segments, similar equations representing their relationship could provide a means of directly modeling a particular variable. The underlying relationships in complex data sets, however, are often nonlinear and limit the usefulness of output models. For example, the linear equation fit to the time series segment data in Figure 9 does a poor job of prediction. A parabolic equation would be better suited to model the relationship between the two variables in this time series segment. As a result, a modeling approach that incorporates multiple variables may offer better insight into the relationships between the variables contained within the data set.

5.1.5.2 Multiple linear regression

Multiple linear regression builds upon the framework initially used by the simple linear regression but extends the concept to build equations that leverage other variables within the data set. More specifically, the multilinear regression model leverages techniques of ordinary least squares to develop equations of the form:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + \epsilon, \tag{1}$$

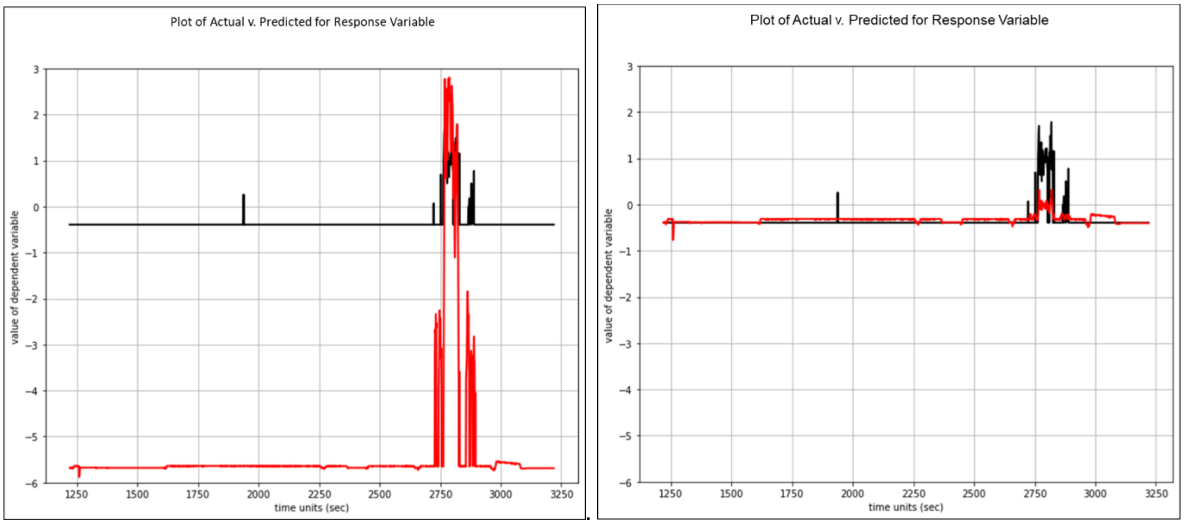
where

y = the dependent variable;
β_i = represent scalars/weights associated with variable x_i, for $i = 1, \ldots, n$;
β_0 = a bias term (value of y when variables have scalars/weights are all zero);
ϵ = the calculated error for the model.

The developed script functionality examines time series segments with outliers, identifying all variables with outlying data points, and generating multivariate linear models using each variable with outlying data points as the dependent variable. Analysis is then performed on each model to determine goodness of fit (Figure 10). If a system variable (or variables) used within the model has a high variance inflation factor (VIF) (i.e., a VIF over 10), the script automates the removal of the variable that decreases mean absolute error (MAE) the most. The results of each model are stored within a data frame (including initial and intermediate models) for later analysis. The resulting models improve the prediction of the dependent variable in comparison to simple linear regression since they incorporate more than one of the other variables within the data set.

In Figure 11, the image on the *left* illustrates an initial multilinear regression predictive model for a particular variable that performs poorly; whereas, in the image on the *right*, the multilinear regression predictive model refined by removing variables with high VIF performs better than the unrefined model.

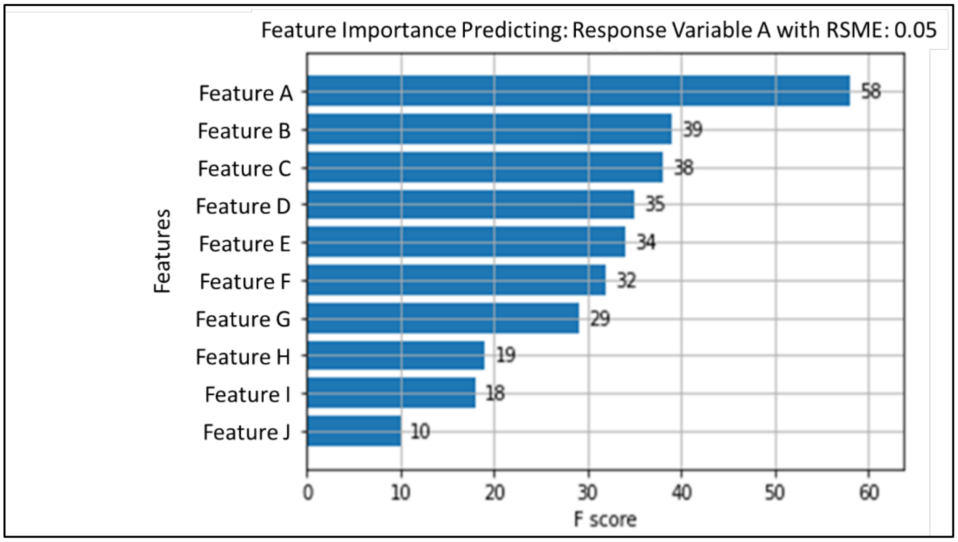
Figure 11. Multi-linear regression model output.



5.1.5.3 Extreme gradient boost (XG Boost)

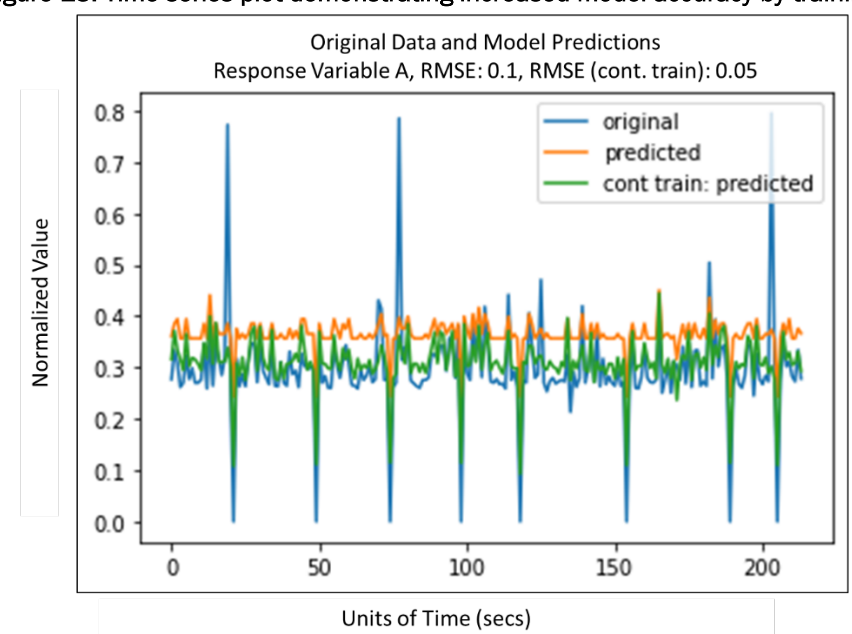
The last method implemented was the extreme gradient boost (XG Boost) (decision tree) method. The XG Boost method examines a data set and branches the decision tree on variables that optimize (in this case, minimize) a loss function. The variables with a greater impact on the loss function are recorded by the script and used to generate an F-score for each variable. These F-scores can then be used to identify predictor variables that have a greater influence on the response variable identified by the decision tree. Figure 12 is an illustrative example of the F-scores of variables within the data set for predicting response variable A. Greater F-scores for model features indicate a greater ability for the model to predict.

Figure 12. F scores for variables to predict response variable A.



By continuously training the model, it is possible to more accurately model and predict a given variable and better capture variables' influence upon a given variable (Figure 13). Resulting models for each data set, as well as the continuously trained model, are output to a file, and plots are generated and automatically stored to directories to expedite future model analysis efforts.

Figure 13. Time series plot demonstrating increased model accuracy by training.



5.2 Linking Reliability Scorecard ratings and reliability outcomes

CCDC DAC collaborators provided AMSAA Reliability Scorecards, which represented an assessment of reliability risk for programs of record pre-Milestone A or as early in program life as possible. CCDC DAC attempted to obtain test data where possible and reliability assessments from DOT&E reports for programs with Reliability Scorecards. Next, CCDC DAC then compared the Reliability Scorecard ratings to program reliability metrics to assess Scorecard correlation with reliability outcomes and the Scorecard's potential usefulness in predicting reliability pre-Milestone A. The Scorecard consists of 8 major categories (a total of 40 subcategories) with weightings based on importance and aggregated to a risk score of 1 to 100, with 1 being the lowest risk and 100 the highest risk. Elements that do not apply because of where the program is in development are not included in the risk score.

If a program receives a high risk assessment for reliability in the Scorecard, it is possible that the vendor and the program office would take extra steps to mitigate this risk, and thus, a high risk score might not be correlated with poor reliability outcomes. In other words, knowledge of this fact might create the conditions for programs to improve. If it turns out that programs with high risk are actually not experiencing more reliability issues, additional research would be needed regarding whether these mitigation steps are taking place, whether programs with low risk are not actually following reliability best practices, or whether reliability issues are endemic regardless of observed reliability best practices.

5.3 Program reliability data crosswalk

5.3.1 AMSAA Scorecards

This portion of the research looked across the different data sources to answer questions such as what systems had AMSAA Scorecards completed? Did these systems also have DOT&E reports? Were cost data available and collected for these systems? If systems were in multiple sources of data, when was the Scorecard evaluation completed (relative to the overall schedule of the program)?

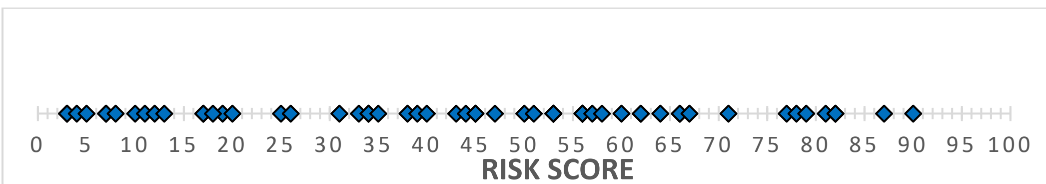
Overall, the AMSAA Scorecards included data of multiple vendors for specific particular program while also containing data for various different system program offices. Therefore, the level of the Scorecard evaluation was not the same throughout. Ultimately, there were limited reliability data found available for system analysis at pre-Milestone A.

When performing the crosswalk with the DOT&E reports, one program with multiple vendors matched the Scorecards and also had reliability performance data in the DOT&E reports. However, of the three vendors, no vendors were selected to continue to the next phase of the acquisition process. Ultimately, another vendor was selected for the program.

For the crosswalk of data with available cost information, two systems were found in both data sets. However, both Scorecards were completed after the Milestone A timeframe.

CCDC DAC analyzed 57 scores from scorecards and the risk results are shown in Figure 14. These risk scores range from very low risk to very high risk, offering wide variability in the predictor variables.

Figure 14. Range of observed Scorecard scores for 57 programs/bids.



5.3.2 Scorecard summary deep dive

Of the available 57 Scorecards, a total of 26 Scorecards were provided for an in-depth review for the program data crosswalk. A review of 26 system CCDC DAC AMSAA Reliability Scorecards were gathered for each of the associated subcategories (Reliability Requirements and Planning, Training and Development, Reliability Analysis, Reliability Testing, Supply Chain Management, Failure Tracking and Reporting, Validation and Verification, and Reliability Improvements). Analyzing these data at the aggregate level revealed that the scores within the categories for Training and Development, and Reliability Improvements average at a high risk across the systems (Table 3). Category scores averaged from 32 to 62. Across the systems, the average scores were generally high or medium risk, with the Failure Tracking and Reporting category and Validation and Verification category scoring lower on average. The one category with consistently high scores was the Training and Development category.

Table 3. Reliability subcategory normalized score summary.

ID #	Overall Risk Assessment	High	Med	Low	Not Eval	Rel Req and Planning	Training and Dev	Reliability Analysis	Reliability Testing	Supply Chain Mgmt	Failure Tracking and	Validation and Verificatio	Reliability Improvements	
1	58	9	23	3	5	64	63	54	50	88	33	50	50	
2	40	9	13	13	5	52	50	12	50	56	17	1	67	
3	82	22	9	1	8	100	100	59	100	50	50	1	100	
4	18	0	16	24	0	7	13	23	29	31	1	30	42	
5	67	17	21	1	1	73	81	65	64	69	50	60	63	
6	46	9	21	9	1	43	63	50	64	50	17	50	38	
7	37	14	3	19	4	17	56	23	36	75	33	40	67	
8	21	5	10	21	4	8	31	12	36	44	1	30	67	
9	51	16	12	12	0	61	63	54	64	56	17	40	33	
10	60	12	24	2	2	55	69	42	57	94	50	90	50	
11	57	7	32	0	1	57	69	50	64	69	50	50	50	
12	47	3	29	7	1	41	38	42	43	69	17	50	50	
13	13	0	10	30	0	11	13	19	14	6	17	1	17	
14	78	20	11	1	8	93	69	69	100	50	100	1	100	
15	44	5	26	8	1	48	38	42	36	56	50	20	50	
16	56	12	22	5	1	41	38	69	50	50	83	50	100	
17	50	5	32	2	1	39	69	50	57	50	50	50	63	
18	32	3	22	14	1	14	25	38	57	56	33	20	50	
19	71	17	13	2	8	89	69	69	90	50	1	1	100	
20	41	6	4	9	21	100	100	46	1	Not Evaluat	17	75	20	
21	53	8	25	4	3	45	69	42	50	100	33	50	50	
22	35	14	2	21	3	53	63	1	1	100	1	20	33	
23	90	34	5	1	0	95	81	92	100	100	67	70	100	
24	43	10	14	11	5	32	88	42	43	50	17	50	67	
25	45	10	14	11	5	27	88	46	50	80	17	50	67	
26	50	6	22	7	5	64	50	42	57	44	33	50	33	
	49.42	10.50	16.73	9.15	3.62	51.12	59.85	44.35	52.42	61.72	32.88	38.46	58.73	Average

5.3.3 Reliability Scorecard links to cost and schedule risk

A review of the prior ground vehicle cost model revealed how the cost model's reliability variable impacted the estimated system LCC. The cost model revealed a lack of sensitivity to the reliability metric. For example, inputting the notional MMBOMF of 600; 6,000; and 12,000 generated different cost outputs (Table 4).

Table 4. Reliability impact on LCC estimate.

MMOBMF	Estimated LCC	Program Acquisition	Operations and Sustainment (O&S)	Disposal
600	$30,708,614	$8,811,179	$21,795,635	$101,800
6,000	$25,725,446	$8,811,179	$16,812,467	$101,800
12,000	$24,933,015	$8,811,179	$16,020,036	$101,800

In this example, the impact of reliability upon the cost is evident in the change in the overall LCC and operations and sustainment cost. Note that both the program acquisition and disposal costs did not change as the MMOBMF changed. In general, the reliability of the system should change the procurement quantity in order to maintain a specified level of availability for the users. For example, a system with a very low reliability (e.g., 600 in this example) should have a significantly higher program acquisition cost in order to maintain a required service level. The number of units procured impacts the number of associated spares to purchase for O&S repairs. Furthermore, the amount required in disposal would also adjust since additional units would require disposal.

5.3.4 Reliability Scorecard linkage to pre-Milestone A cost analysis

The list of systems included in the Scorecard data were reviewed against the list of systems collected in the related cost research. From the analysis, it appears that only one system within the cost research had a Scorecard completed. However, the subject system entered Milestone A in December 2011 and Milestone B in August 2015. The date of the AMSAA Scorecard completion appears to be after completion of Milestone A. Between the two systems under evaluation for the program, one system had a risk score of 32 (lower risk), while the other had a risk score of 50 (higher risk).

5.3.5 Ground vehicle data on HPC

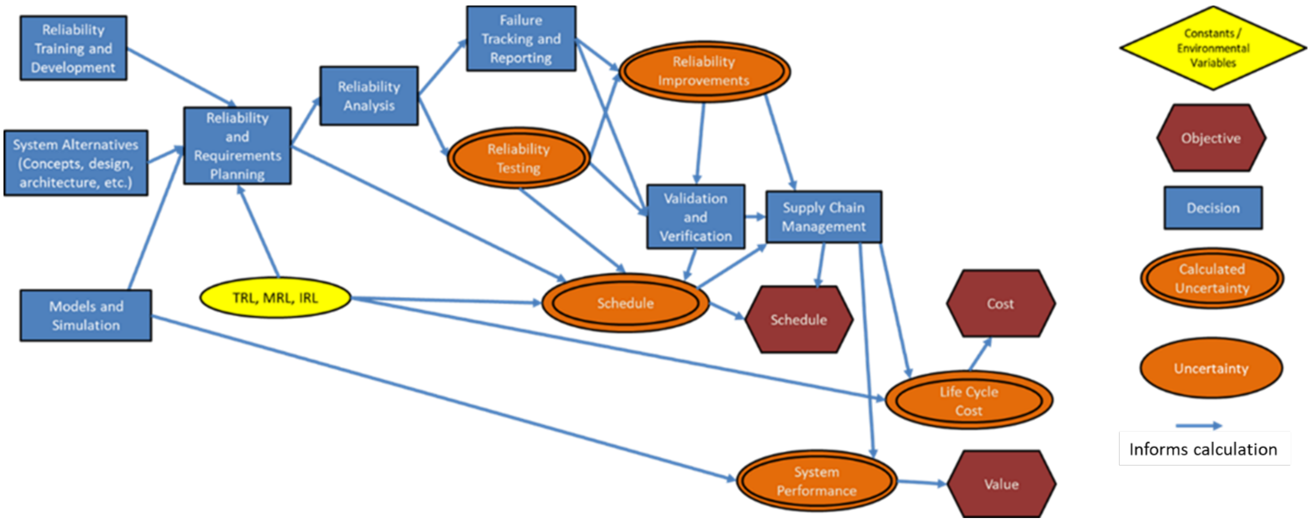
A review of the 42 systems with data on the HPC revealed that only one system received analysis via the early Reliability Scorecard. This system received both an initial assessment as well as a final assessment. This system was an ACAT II system, and the Scorecard was completed in the 2014 timeframe. Between the timeframe of the two assessments, the system initially had a risk score in the high 60s and then reduced to a score slightly below 50 to indicate a medium risk. The system improved primarily in the Failure Tracking and Reliability Improvements categories. However, the scores in the other categories remained near or above 50 on most other categories. The system also was mentioned in DOT&E reports as having "Poor system reliability." However, the metric of the reliability performance (e.g., Mean Miles Between Mission Aborts) was not provided in the DOT&E reports. Also, this system moved into production, which indicates that the Scorecard was not completed pre-Milestone A. Finally, none of the systems were included in the available LCC analysis data set.

5.4 Reliability program value model

5.4.1 Influence diagram

An influence diagram (Figure 15) was developed in support of generating a value model, which leveraged the eight critical reliability categories outlined in the AMSAA Scorecard. Each category was integrated with an existing influence diagram to illustrate additional factors related to reliability. The updated influence diagram shows how each of the Scorecard elements could relate to schedule, cost, and value.

Figure 15. Reliability program influence diagram.



5.4.2 Value model

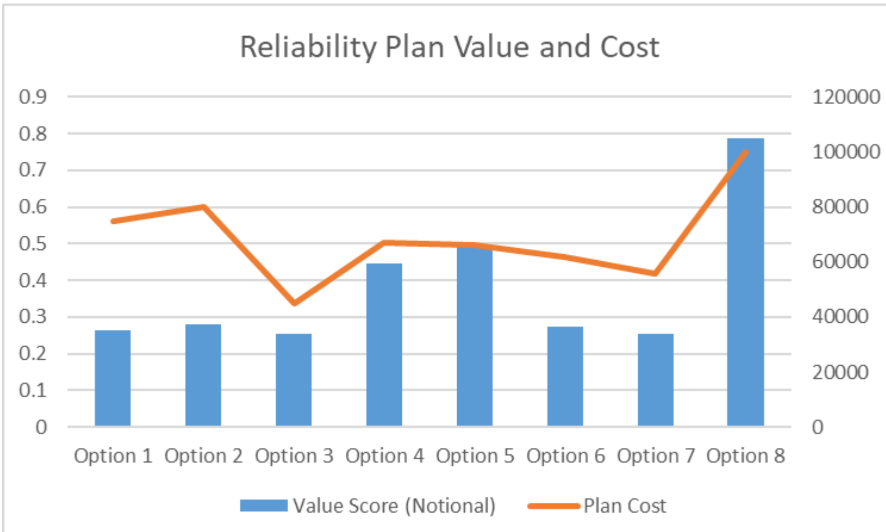
A notional value model construct (Figure 16) depicts how elements from the AMSAA Scorecard could be leveraged to develop a value model as well as the trade-off graph between the reliability plan value and cost. This example model uses categories and metrics of information that could be identified at pre-Milestone A and that could be quantified by the program office, who could determine which elements and measures are most important. These selections could be utilized in the value model framework to illustrate the trade-off between reliability program cost investments and value. It could also be used to show the reliability program cost investment versus the impact on the LCC and the program schedule.

Figure 16. Reliability program value model notional construct.

Value Model							
Objectives	Identify and incorporate reliability activities into the program Integrated Master Schedule	Increase reliability engineering staff incorporation with design team	Improve Programmatic Coverage through Reliability Analysis	Identify and mitigate failure modes with Reliability Testing	LCC	Schedule
Measures	% of reliability tests integrated into IMS	% of reliability team incorporated with design team	% of system with failure modes identified	% of reliability growth plan completed		
Threshold	60%	10%	25%	20%	Higher Impact	Higher Impact
Objective	100%	90%	80%	100%	Lower Impact	Lower Impact

The notional objectives were selected from a subcategory within the AMSAA Scorecard. The threshold and objective values are notional and could be defined by a specific program office. Figure 17 shows a comparison output of a variety of different reliability plan value scores (e.g., Option 1 achieves a value of 25% and costs approximately $78,000, while Option 5 achieves a value of 50% and costs $70,000).

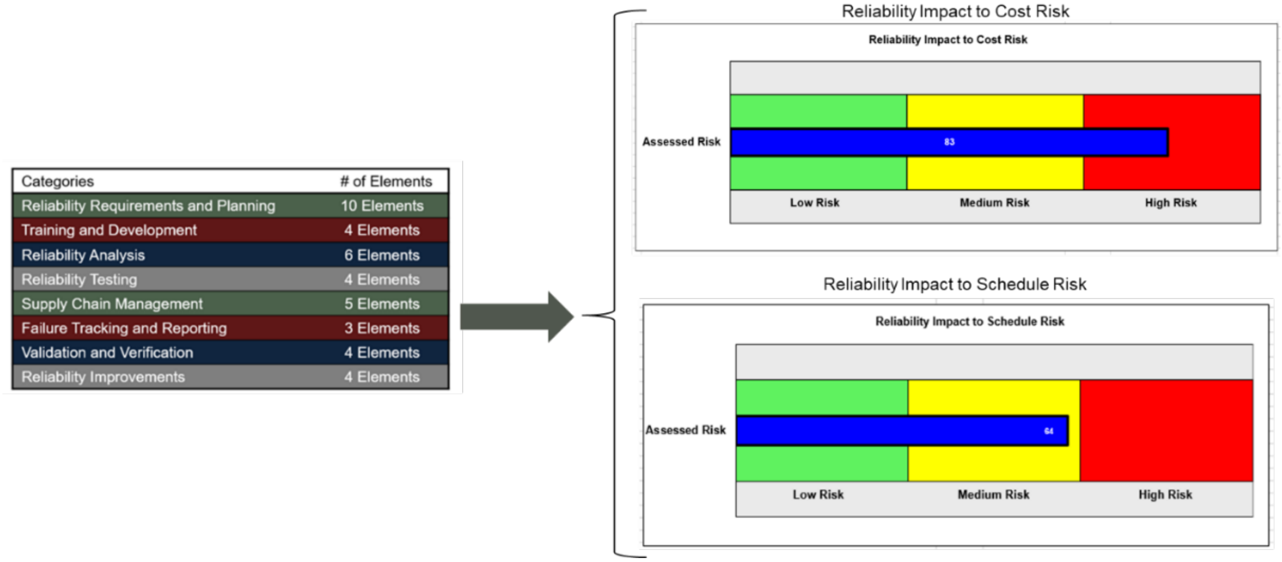
Figure 17. Reliability program value and cost



5.4.3 Reliability Scorecard links to cost and schedule risk

In order to illustrate the linkage between reliability, schedule, and cost, additional output for the AMSAA Reliability Scorecard was created. Currently, the Scorecard outputs a single general risk score. Using the existing Scorecard's 8 categories and 40 elements, each element was assigned to impact either cost or schedule risk. Then, using the same weighting and methodology already within the Scorecard, a new risk score output was generated that linked the reliability impact to cost and the reliability impact to schedule (Figure 18). This update could help program managers better understand how the reliability decisions and factors might impact the cost and the schedule risk.

Figure 18. New reliability scorecard with impact to cost and schedule risk.

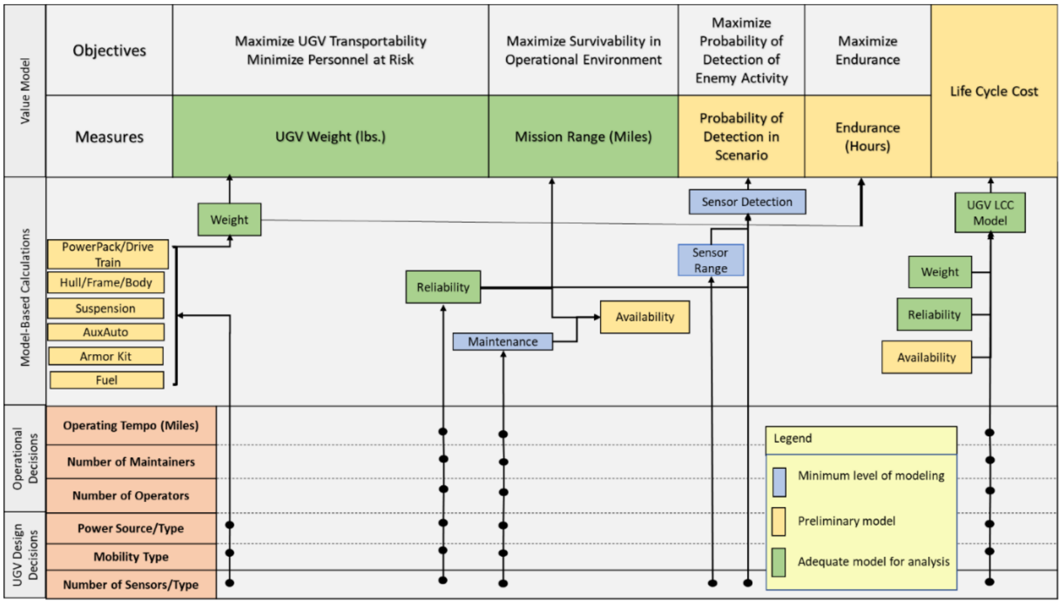


5.5 Integrated modeling framework

The research team created an assessment flow diagram (AFD), shown in (Figure 19), to illustrate the flow of information and the models used to calculate performance measures and the life cycle cost for a UGV. The AFD starts with system design decisions such as mobility, power source, and sensor technology capabilities. Design and operational decisions are inputs to the models shown in the model-based calculations section, which determine the system alternatives' performance measures and life cycle cost. The color-coding in describes the current modeling progress. In the legend, the blue color represents future additions to the UGV model and were implemented in this case study. The yellow color indicates the use of lower fidelity models. The green color indicates the use of higher fidelity models.

UGV notional objectives include maximizing transportability, survivability in operational environments, probability of enemy detection, and endurance. Performance measures are developed for each objective and, in the UGV model, include total vehicle weight, mission range, probability of detection, and endurance. The impact of design decisions on reliability, performance, and life cycle cost are displayed.

Figure 19. Assessment flow diagram.

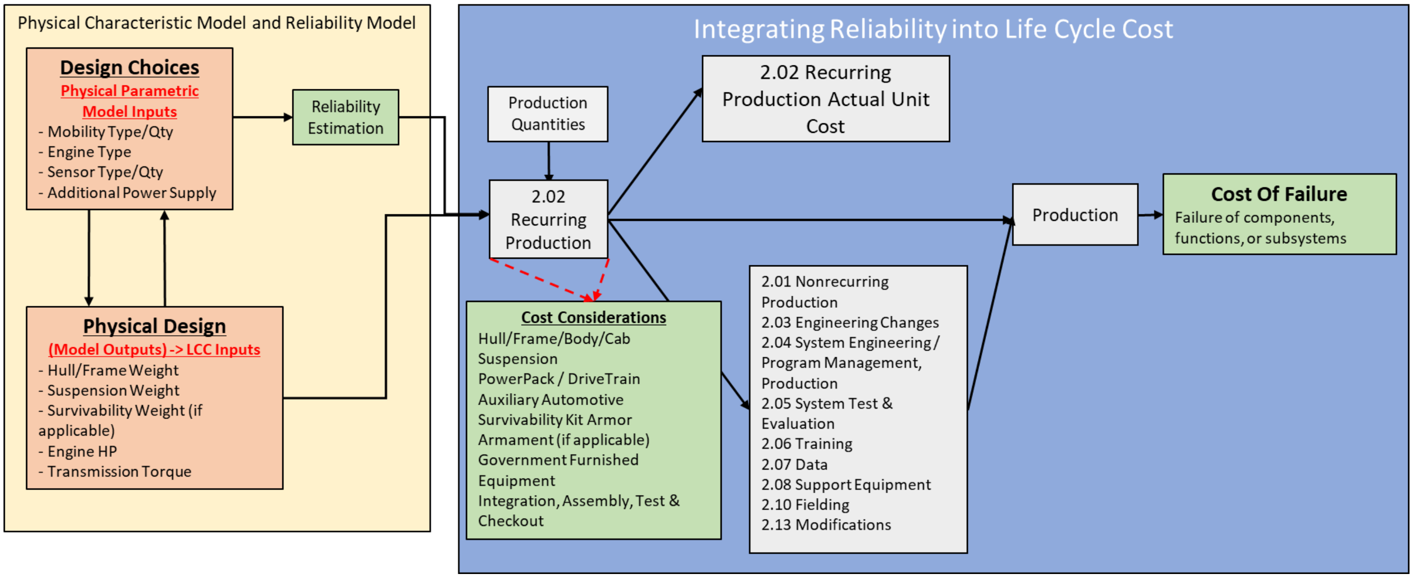


5.5.1 Adapting the ground vehicle cost model for an unmanned system

Previous efforts created a cost model for wheeled ground vehicles for the previous cost-related research.* This cost model was adapted for the UGV use case, and a mapping of the cost model is shown in Figure 20. This diagram provides a high-level view of procurement costs. The figure shows how design decisions are used as inputs into the cost model and how the cost model calculates LCC from those decisions. The research team focused on adapting the cost model to emphasize reliability in recurring production and adding failure cost components to the LCC. This process added an element of traceability from design decisions, physical performance, and reliability estimates of system functions or subsystems to the cost of failure in the LCC, which did not previously exist. Other considerations of restructuring focused on how the LCC changes from a manned to an unmanned platform and the change of personnel and management costs.

* Jeff Cherwonik, 2017, "Engineered Resilient Systems (ERS) Life-Cycle Cost Analysis for Trade-Space Generation," unpublished report, Vicksburg, MS.

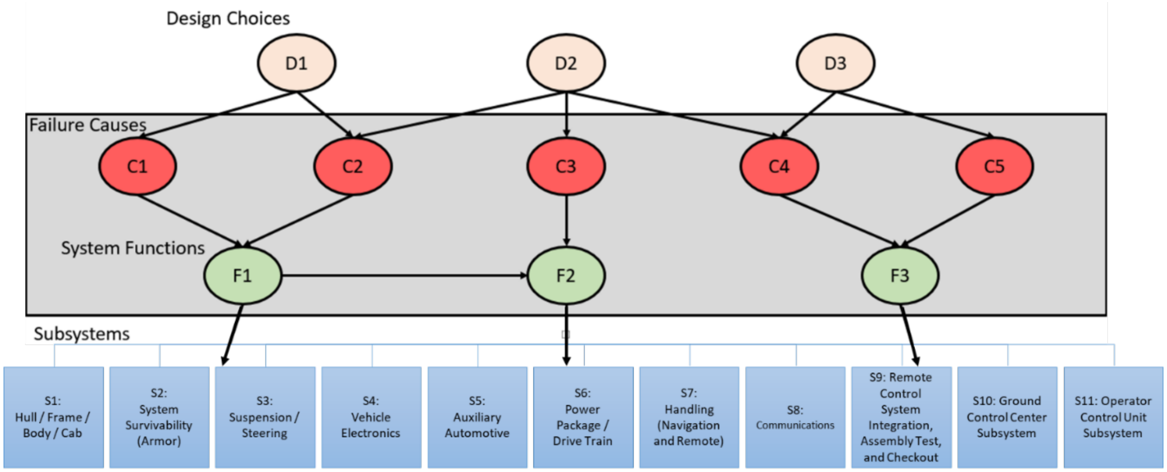
Figure 20. Adaptation of the ground vehicle cost model for an unmanned system.



5.5.2 Function mapping to Bayesian network

Bayesian networks are networks of connected variables that generate predictions based on correlations and assumptions. Generally, Bayesian networks are used when there is a lack of data available or when analysts evaluate an event that occurred and predict the likelihood that any one of several possible known causes was the contributing factor. The Bayesian network architecture uses the elements shaded in the *gray area* in Figure 21. The figure shows that the functional analysis ties the design decisions to the subsystem- and system-level reliability estimate using functional block diagrams and Bayesian networks. Program managers could make design choices that are used as inputs into the Bayesian network. The goal is to estimate failure propagation throughout the functions that can directly or indirectly impact subsystem reliability.

Figure 21. Function mapping to Bayesian networks.



6 Conclusions

Gaining access to DoD system reliability data is difficult across the life cycle as evidenced from the program crosswalk investigation as well as the AMSAA Scorecard research. For the ground vehicle data, the lack of data consistency within data sets presented obstacles to analysis that required additional time for data cleaning and model development. DoD systems' reliability data are varied across the life cycle. The information varies from subjective information, actual test results, and actual performance data. However, no systems examined contained reliability information that could be compared across the explored data sets available. The lack of availability of program data during pre-Milestone A analysis creates difficulty for developing accurate linkages between reliability and cost impacts.

For this research, the team acquired relevant data sets and applied new techniques to early life-cycle reliability analysis, including XG Boost regression with time series segmentation to predict reliability metrics. This research also created a hybrid model composed of the in situ and post hoc approaches that have predictive reliability capabilities, leveraged the HPC to obtain model results for large data sets, and identified the relative importance of variables that can be used for reliability prediction. Furthermore, methods for qualitatively evaluating reliability programs were created through a value model based on Reliability Scorecards and available data at pre-Milestone A. In order to highlight the linkages of reliability to cost, this research implemented schedule and cost risk output for the AMSAA Scorecard. Finally, the research successfully created an integrated model that included design, reliability, performance, value, and cost.

7 Future Work

Suggestions for future work include using this data analysis for ground vehicles to inform designs of future vehicles. Furthermore, Scorecard analysis could be performed across the DoD to allow for more informed programmatic decisions for other systems. This information could potentially improve more programs and would allow for refined Scorecard weights. Finally, another area of additional research includes further investigation of the links between cost and reliability in established models. This includes building upon the initial framework of the value model, investigating the fidelity of linkages between the cost and schedule risk of the new scorecard output, and integrating additional reliability elements into the LCC ground vehicle cost model.

Bibliography

Bond, William Glenn, Haley Dozier, Thomas L. Arnold, Michael Y. Lam, Quyen T. Dong, Indu Shukla, Brandon E. Hansen, Antoinette R. Silas, Jeremiah Prieto, and Corey Hollis Mize. 2020. "A Hybrid Learning Approach to Prognostics and Health Management Applied to Military Ground Vehicles Using Time-Series and Maintenance Event Data." In *Annual Conference of the PHM Society* 12, no. 1 (November): 1–10. https://doi.org/10.36001/phmconf.2020.v12i1.1146.

Brady, Shawn P. 2013. "A Reliability Approach to Risk Assessment and Analysis Alternatives: Based on a Ground Vehicle Example." *Defense Acquisition Technology and Logistics*, July–August 2013, 28–33. Defense AT & L Magazine(dau.edu).

Defense Acquisition University. "CCDC Reliability Scorecard", Accessed September 8, 2022. https://www.dau.edu/tools/t/CCDC-Reliability-Scorecard#:~:text=The%20CCDC%20Reliability%20Scorecard%20examines,of%20the%20overall%20Reliability%20Program.

Defense Acquisition University. "Reliability Key System Attribute (KSA)", Accessed August 21, 2022. https://www.dau.edu/acquipedia/pages/ArticleContent.aspx?itemid=372 accessed 08212022.

DoD (Department of Defense). 2005. *DoD Guide for Achieving Reliability, Availability, and Maintainability*. Washington, DC: Department of Defense. https://www.acqnotes.com/Attachments/DoD%20Reliability%20Availability%20and%20Maintainability%20(RAM)%20Guide.pdf.

Fiondella, Lance, and Panlop Zeephongsekul. 2011. "Reliability of Systems with Identically Distributed Correlated Components." In *Proceedings—17th ISSAT International Conference on Reliability and Quality in Design*, 4 August 2011–6 August 2011, Vancouver, Canada, 26–30.

Fiondella, Lance, and Luidong Xing. 2014. "Reliability of Two Failure Mode System Subject to Correlated Failures." In *Proceedings—Annual Reliability and Maintainability Symposium*, 27–30 January 2014, Colorado Springs, CO, 1–6. https://doi.org/10.1109/RAMS.2014.6798438.

Hu, Zhengwei, and Xiaoping Du. 2018. "Integration of Statistics- and Physics-Based Methods—A Feasibility Study on Accurate System Reliability Prediction." *Journal of Mechanical Design* 140, no. 7 (July): 074501. https://doi.org/10.1115/1.4039770.

Littlefield, Scott, Thomas Mazzuchi, and Shahram Sarkani. 2012. "Predicting Reliability in Design of Complex Systems with Common-Cause Failures and Time-Varying Failure Rates." In *2012 IEEE International Systems Conference SysCon 2012*, 19–22 March 2012, Vancouver, Canada, 1–5. https://doi.org/10.1109/SysCon.2012.6189455.

Nguyen, Huy, and Rong Zheng. 2013. "A Binary Independent Component Analysis Approach to Tree Topology Inference." *IEEE Transactions on Signal Processing* 61, no. 12 (June): 3071-3080. https://doi.org/10.1109/TSP.2013.2254476.

Office of the Secretary of Defense in collaboration with the Joint Staff. 2009. *Department of Defense Reliability, Availability, Maintainability, and Cost Rationale Report Manual*. Washington, DC: Office of the Secretary of Defense. https://www.acqnotes.com/Attachments/The%20Reliability,%20Availability,%20Maintainability%20&%20Cost%20(RAM-C)%20Rationale%20Report%20Manual.pdf.

OUSD(A&S) (Office of the Under Secretary of Defense for Acquisition and Sustainment). 2015. *Operation of the Adaptive Acquisition Framework*. Department of Defense Instruction 5000.02, 23 January 2020. Washington, DC: Department of Defense. https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodi/500002p.pdf.

Thodi, Premkumar, Faisal Khan, and Mahmoud Haddara. 2013. "Risk Based Integrity Modeling of Offshore Process Components Suffering Stochastic Degradation." *Journal of Quality in Maintenance Engineering* 19, no. 2 (May): 157–180. https://doi.org/10.1108/13552511311315968.

The Office of the Director, Operational Test and Evaluation. n.d. "DOT&E Annual Reports," Accessed November 4, 2020. https://www.dote.osd.mil/annualreport/.

Abbreviations

Abbreviation	Term
ACQ	Acquisition Costs
AFD	Assessment Flow Diagram
AMSAA	Army Materiel Systems Analysis Activity
CAN	Controller Area Network
CCDC	Combat Capabilities Development Command
DAC	Data & Analysis Center
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
DoD	Department of Defense
DOT&E	Director, Operational Test and Evaluation
ERDC	Engineer Research and Development Center
FEA	Finite Element Analysis
FMEA	Failure Modes and Effects Analysis
FTA	Fault Tree Analysis
JCIDS	Joint Capabilities Integration and Development System
KPP	Key Performance Parameter
KSA	Key System Attribute
LCC	Life-cycle cost
MAE	Mean Absolute Error
MDT	Mean Down Time
MMBOMF	Mean Miles Between Operational Mission Failure

MTBF	Mean Time Between Failures
MTFR	Mean Time Failure Rate
MTTM	Mean Time to Maintenance
MTTR	Mean Time to Repair
NaN	Not a Number
O&M	Operations and Maintenance
O&S	Operations and Sustainment
PM	Project Managers
RAM	Reliability, Availability, and Maintainability
R&D	Research and Development
R&M	Reliability and Maintainability
SBD	Set-Based Design
tSNE	T-distributed Stochastic Neighbor Embedding
UARK	University of Arkansas
UMAP	Uniform Manifold Approximation and Projection
VIF	Variance Inflation Factor
VIN	Vehicle Identification Number
XG Boost	Extreme Gradient Boosting

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 instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this 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 Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
December 2022	Final	

4. TITLE AND SUBTITLE
Early Life-Cycle Prediction of Reliability

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT
0603463A

6. AUTHOR(S)
Randy K. Buchanan, Christina H. Rinaudo, George E. Gallarno, and Matieu L. Lagarde

5d. PROJECT NUMBER
BP4

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
US Army Engineer Research and Development Center (ERDC)
Information Technology Laboratory (ITL)
Waterways Experiment Station, 3909 Halls Ferry Road
Vicksburg, MS 39180-6199

8. PERFORMING ORGANIZATION REPORT NUMBER
ERDC/ITL TR-22-1

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Headquarters, US Army Corps of Engineers
Washington, DC 20314-1000

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Distribution Statement A. Approved for public release: distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The intent of this project is to investigate a variety of approaches for the development of a basic model for the early life-cycle prediction of reliability (pre-Milestone A). The United States Department of Defense (DoD) currently utilizes an acquisition framework in which system development advances through a series of checkpoints known as milestones. Each milestone represents a decision point, with Milestone A being the earliest in the life cycle. At Milestone A, also known as the risk-reduction decision, the DoD evaluates design concepts while also committing funds to the maturation of technologies in an effort to mitigate future risks. Typically, little is known about the particular system to be developed at this point in the acquisition life cycle, but DoD regulations require program managers to submit system reliability information (OUSD[A&S] 2015). Traditional reliability predictions, however, require extensive knowledge of the system of interest to produce accurate results. This level of knowledge is unavailable at or before Milestone A, therefore, there is a need to create models and methodologies for the prediction of system reliability. This report provides an overview of a variety of methods investigated to improve the prediction of early life cycle reliability.

15. SUBJECT TERMS
Systems engineering; Engineering design; Military research; Reliability (Engineering)—Evaluation; Reliability (Engineering)--Forecasting

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER (include area code)
Unclassified	Unclassified	Unclassified	SAR	52	