ABSTRACT

Title of Dissertation: BAYESIAN METHODOLOGY FOR RELIABILITY GROWTH PLANNING AND PROJECTION FOR DISCRETE-USE SYSTEMS UTILIZING MULTI-SOURCE DATA

Paul John Nation, Doctor of Philosophy, 2021

Dissertation directed by: Professor Mohammad Modarres, Department of Mechanical Engineering

This research aims to present a Bayesian model for reliability growth projection and planning for discrete-use systems suitable for use throughout all stages of system development. Traditional discrete-use models for reliability growth utilize test data from individual test events at the current stage of development. They often neglect the inclusion of historical information from previous tests, testing similar systems, or eliciting expert opinion. Examining and using data attained from prior bench analyses, sub-system tests, or user trial events often fails to occur or is conducted poorly. Additionally, no current approach permits the probabilistic treatment of the initial system reliability at the commencement of the test program in conjunction with the management variables that may change throughout the execution of the test plan. This research aims to contribute to the literature in several ways. Firstly, a new Bayesian model is developed from first principles, which considers the uncertainty

surrounding discrete-use systems under arbitrary corrective action regimes to address failure modes. This differs from current models that fail to address the randomized times that corrective actions to observed failure modes may be implemented depending on the selected management strategy. Some current models only utilize the first observed failure on test, meaning a significant loss of information transpires if subsequent failures are ignored. Additionally, the proposed strategy permits a probabilistic assessment of the test program, accounting for uncertainty in several management variables.

The second contribution seeks to extend the Bayesian discrete-use system projection model by considering aspects of developmental, acceptance, and operational testing to formulate a holistic reliability growth plan framework that extends over the entire system lifecycle. The proposed approach considers the posterior distribution from each phase of reliability growth testing as the prior for the following growth test event. The same methodology is then employed using the posterior from the final phase of reliability growth testing as the prior for acceptance testing. It then follows that the acceptance testing posterior distribution forms the prior for subsequent operational testing through a Bayesian learning method. The approach reduces unrealistic and unattainable reliability demonstration testing that may result from a purely statistical analysis. The proposed methodology also permits planning for combined developmental and acceptance test activities within a financially constrained context.

Finally, the research seeks to define an approach to effectively communicate developmental system reliability growth plans and risks to decision makers. Like many of their other specialist science peers, reliability professionals are fantastic communicators – with other reliability practitioners. However, when reliability professionals move beyond their world to make an impact, they often face the same challenge scientists from every discipline face – the difficulties of clearly communicating science to their audience.

BAYESIAN METHODOLOGY FOR RELIABILITY GROWTH PLANNING AND PROJECTION FOR DISCRETE-USE SYSTEMS UTILIZING MULTI-SOURCE DATA

by

Paul John Nation

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2021

Advisory Committee:
Professor Mohammad Modarres, Chair
Professor Aris Christou
Assistant Professor Katrina Groth
Professor Jeffrey Herrmann
Doctor Martin Wayne (Special Advisor)
Professor Gregory Baecher (Dean's Representative)

Preface

This work was borne of my desire to do more, to be more, and to explore. I never expected my life or career to get to where it is now, but I am certainly glad it did.

I have discovered how reliability growth planning models and methodologies can be improved throughout my research journey, coupled with my typical day-to-day career. More importantly, I have become enlightened as to why they need to improve. Quite simply, if organizations do not change how reliability growth plans are developed, they will continue to waste resources and fail to achieve optimal results. In my current role, I routinely observe reliability growth plans go awry. Some projects fail to invest the required resources. In contrast, others attempt to 'test in' the required reliability without a realistic understanding of how reliability can be influenced by management strategy, the effectiveness of corrective actions and changes that are designed to address failure modes, or the initial system reliability upon entering the program. Furthermore, each of these reliability growth program aspects is grounded in a level of uncertainty. Management strategies drift as constraints and personalities change. The effectiveness of corrective actions can never be sure without significant testing and use, which may be well outside of the reliability program's scope. Moreover, initial reliability is never known for sure; it is only ever an estimate within confidence bounds.

The basis of this research initially stemmed from my passion for improving the reliability of developmental systems for military applications based on my observations and negative experiences. As military acquisition budgets alter over time, reliability practitioners are often asked two specific questions: How can we do more with fewer resources, and how can we maximize the use (and often reuse) of data that is expensive to attain? My passion is to find out and develop the tools and processes necessary to make it easier for future engineers.

Dedication

For my family,

you were supportive despite the long nights and coffee; I love you with all my heart

&

For my parents,

thank you for believing in me; I hope I have made you proud.

Acknowledgements

I want to express my thanks and gratitude to my supervisor, Professor Mohammad Modarres, for his positive and supportive guidance. You have provided me with knowledge, experience, and insight. I thank you for your patience and dedication.

This work would undoubtedly be incomplete without the assistance of Doctor Martin Wayne. You have acted as a mentor, a sounding board for my ideas, and provided me with significant guidance. You have inspired me to seek new knowledge, and for that, I will always be grateful.

Thank you also to the members of my Dissertation Committee for your time, efforts, and valuable contributions to your various fields of endeavor.

Last but certainly not least, I want to thank my family, friends, and work colleagues for their support, encouragement, and feedback – it has not gone unnoticed and is appreciated.

"Learning is the beginning of wealth. Searching and learning is where the miracle process all begins. The great breakthrough in your life comes when you realize it that you can learn anything you need to learn to accomplish any goal that you set for yourself. This means there are no limits on what you can be, have or do."

~ Albert Einstein ~

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AMPM	AMSAA Maturity Projection Model
AMSAA	U.S. Army Materiel Support Analysis Activity (past terminology)
AT	Acceptance Testing
AT&E	Acceptance Test and Evaluation
CAP(s)	Corrective Action Period(s)
CI	Confidence Interval
DT	Developmental Testing
DT&E	Developmental Test and Evaluation
FEF	Fix Effectiveness Factor
FMEA	Failure Modes and Effects Analysis
FMECA	Failure Modes, Effects and Criticality Analysis
HPP	Homogeneous Poisson Process
LCB	Lower Confidence Bound
LUT	Limited User Test
M&S	Modeling and Simulation
MS	Management Strategy
MTBF	Mean Time Between Failure
MTBFg	MTBF Goal
$MTBF_i$	MTBF Initial
$MTBF_R$	MTBF Requirement
NHPP	Non-Homogeneous Poisson Process
OC	Operating Characteristic (curve)
OT	Operational Testing

OT&E	Operational Test and Evaluation
PM2	Planning Model Based on Projection Methodology
PM2-C	Planning Model Based on Projection Methodology - Continuous
PM2-D	Planning Model Based on Projection Methodology - Discrete
RCM	Reliability Centered Maintenance
RDT	Reliability Demonstration Test
RGC	Reliability Growth Curve
RGP	Reliability Growth Potential
RGT	Reliability Growth Test
ROCOF	Rate of Occurrence of Failures
RPM	Reliability Planning and Management
SPLAN	AMSAA System Level Planning Model
SSPLAN	AMSAA Subsystem Level Planning Model
T&E	Test and Evaluation
TAFT	Test-Analyze-Fix-Test Process
UCB	Upper Confidence Bound

Mathematical Notation

λ_{AT}	Acceptance test failure intensity
M_{AT}	Acceptance test mean time between failure
t	Arbitrary time
α	Consumer's (user's) risk
λ_{DT}	Developmental test failure intensity
M_{DT}	Developmental test mean time between failure
E, ε	Expectation value
λ_{A}	Failure intensity due to A-modes
λ_{B}	Failure intensity due to B-modes
λ_i	Failure intensity mode *i*
μ_d	Fix effectiveness factor
a_μ	Fix effectiveness factor Beta distribution alpha parameter
b_μ	Fix effectiveness factor Beta distribution Beta parameter
a	General Beta distribution alpha parameter
b	General Beta distribution Beta parameter
β	General gamma distribution scale parameter
α	General gamma distribution shape parameter
λ_G	Goal failure intensity
M_G	Goal mean time between failure
λ_I	Initial failure intensity
M_I	Initial mean time between failure
MS	Management strategy
a_m	Management strategy Beta distribution alpha parameter

b_m	Management strategy Beta distribution Beta parameter
M_R	Mean time between failure requirement
λ_{OT}	Operational test failure intensity
M_{OT}	Operational test mean time between failure
β	Producer's (contractor's or developer's) risk
β	Reliability growth curve shape paramete
T_{DT}	Total developmental test time
T_{RDT}, T_{AT}	Total reliability demonstration test time (acceptance test time)
$\tilde{\beta}$	Underlying failure gamma distribution scale parameter
$\tilde{\alpha}$	Underlying failure gamma distribution shape parameter
Var	Variance

Chapter 1: Introduction and Motivation

1.1 *Background and Motivation*

Reliability growth is not a new concept within engineering or statistics and first emerged in the early 1960s [1]. In both an industrial and a military context, reliability growth during system development has been recognized as essential for decades, and it continues to evolve substantially. Unfortunately, its power, utility, and beneficial application in many current efforts remain widely unrecognized.

Generally speaking, the first prototypes or items produced during the development of a new complex system will contain design, manufacturing, or engineering deficiencies. Because of these deficiencies, the initial reliability of the newly developed system may be below the reliability goal or requirement. In order to identify and correct these deficiencies, items under design and development are often subjected to a rigorous reliability testing program. During testing, problem areas are identified, and appropriate corrective actions may be taken to improve reliability. Thus, reliability growth is the improvement in the reliability of a product, component, subsystem, or system over a period of time due to changes in the design, the manufacturing process, or both [2].

The concept of reliability growth is not simply hypothetical. The reliability growth rate is related to factors such as the management strategy toward taking corrective actions, the effectiveness of the fixes implemented, reliability requirements, the initial system reliability, reliability funding, and business competitive factors.

The management strategy may be driven by budget and schedule; however, it is defined by management's actual activities and decisions in correcting reliability problems. If the reliability of a failure mode is known through analysis or testing, then management makes the decision either not to fix (take no corrective action) or to fix that failure mode (implement a corrective action) [3]. Generally, if the reliability prediction by avoiding the failure mode meets management expectations, then no corrective actions would be expected. On the other hand, if the reliability is below expectations, the management strategy would typically call for implementing some form of corrective action. Different management strategies may lead to different reliability outcomes from the same baseline design. For example, one management team may take corrective actions for 80% of the failure modes that surface during reliability testing. In contrast, another management team with the same design and test information may implement corrective actions on only 60%.

The effectiveness of the corrective actions must also be considered. A corrective action typically does not eliminate a failure mode and prevent it from recurring; instead, it reduces its occurrence rate [2]. A corrective action for a problem failure mode may remove a failure altogether or, more frequently, removes a certain amount of its failure intensity. Therefore, a level of intensity will likely remain in the system. The fractional decrease in the problem mode failure intensity due to the corrective action is known as the fix effectiveness factor (or FEF). The FEF will vary from failure mode to failure mode and between various system types, but a typical average for complex military systems has been reported [4] to be about 0.70. With an FEF equal to 0.70, a corrective action for a failure mode is considered to remove 70% of the failure intensity, but 30% remains. As a result, the failure mode may still surface in use or testing; however, its failure intensity has reduced.

The effectiveness of the corrective actions is also relative when compared to the initial reliability at the beginning of testing. Suppose corrective actions give a three-fold improvement in reliability for equipment that initially had one-tenth of the reliability goal. In that case, this is not as significant as a 50% improvement in reliability if the initial system reliability was half the reliability goal.

An effective and efficient reliability growth test planning and management strategy can contribute significantly to developing any new product or system. A solid plan gives the design or development team the focus and ability to meet desired reliability goals on schedule and within project budget constraints. For project management professionals, a well-considered and structured reliability growth strategy permits allocating resources to achieve the desired reliability outcomes using a measured and well-judged approach. Suppose the project management team has an intimate understanding of planned reliability targets. In that case, the reallocation or addition of unnecessary resources can be deliberated depending on reliability growth results.

Military decision making and risk management have always been conducted in a very systematic way by many organizations. The laws and procedures that govern military decisions are typically presented as a foundational policy that is well documented and constitutes a "traditional" guiding principle. One of the downsides to such a rigid and structured approach to commonplace activities is that the doctrinal approaches developed offer little flexibility and takes decades to change. Military systems reliability growth planning and management is no different and is guided through the use of handbooks, policy statements, process publications, and procedures used domestically by other service branches or internationally by other militaries.

The key defense reliability growth publication used by the United States Department of Defense is Military Handbook 189C (MIL-HDBK-189C) Reliability Growth Management which was first developed in February 1981 [5]. Since the original publication, several revisions have been released, the latest being June 2011 [6] [7]. Within many western militaries, this publication is regarded as the authoritative publication relating to reliability growth. The reliability growth concepts and methodologies presented in the publication have been developed over the past few decades with actual application to Army, Navy, and Air Force systems. The evolution of individual tool applications has been developed to the point where considerable payoffs in system reliability improvement and cost reduction can be achieved [7]. However, strict adherence to any set of procedures or guidelines in a narrow-focused manner is not recommended and is likely to have shortcomings. It is important to note that MIL-HDBK-189C is not intended to be relied upon to produce a reliability growth plan without tailoring. Such tailoring is only possible with knowledge of the system in question and its developmental program.

When developing a reliability growth program, numerous uncertainties exist, making planning and the allocation of appropriate resources difficult. It often results in finite resource wastage as allocation is attributed to areas where the most significant impact is not realized. The three critical areas of uncertainty, which should be included in any mathematical model to ensure an accurate reliability growth plan that is suitably resourced to permit successful outcomes, include:

- the management strategy (or MS),
- the FEF, and
- the initial reliability of the system undergoing test upon entering the growth program [8].

Many of the current statistical modeling methods used to develop a sound reliability growth program outlined in MIL-HDBK-189C fail to consider potential sources of uncertainty. In essence, they simply result in a plotted steady reliability growth outcome that extends from program entry to its conclusion that is not reality balanced. Many of these plans and programs refer to “idealized” reliability growth when in fact, reliability improvement hinges mainly on a collection of unknowns. Failing to consider uncertainties often results in a reliability growth plan that displays little resemblance to actual system test performance with marginal reliability improvements and significant resource wastage consequences. In reality, the singular growth curve developed from strict mathematical conformance does not represent the infinite number of paths that reliability growth can take. The singular curve leads to the belief that if the tested system’s reliability followed the provided idealized growth curve, then the system is on track to meeting the stipulated reliability requirement.

Nevertheless, what if the tested system deviates from the idealized growth curve as it will almost inevitably do in practice? How far above or below the idealized curve can the tested system be before the project management team can declare either that the system is on track for early success or that the system is displaying evidence that it will likely not meet the reliability requirement? This revelation leads to an additional underlying reliability communication issue. How can reliability professionals communicate actual reliability performance against planned reliability growth in a standardized manner that considers planning uncertainties and uncertainties in the reliability the system demonstrates during test and development?

By considering the unknowns during both planning and testing, it is possible to better model reliability growth and produce a reliability growth plan that provides the project management team greater clarity when considering reliability performance against stipulated requirements at any given time during test activities.

Many of the procedures outlined in MIL-HDBK-189C are based on a statistical appreciation of reliability growth modeling. Statistical reliability currently rests at an interesting juncture. On the one hand, reducing budgets in areas such as the acquisition of systems requires organizations to “do more with less” and “use (and reuse) all data available.” On the other hand, the era of “big data” means that information from sensors, field failures, and warranty claims can supplement existing reliability growth data, enhancing knowledge and reducing uncertainty. These challenges drive the need to combine multiple information elements from varied sources using models that account for differences in the variation and uncertainty in the sources. Bayesian approach methodologies provide a natural, intrinsic and structured way to combine this information flexibly.

1.2 *Research objectives*

Systems are generally defined by their usage. Continuous use systems have their usage recorded on a continuous scale; examples include hours operated or miles driven. Discrete-use systems operate via discrete user demands. Examples of discrete-use systems include explosive ordnance, such as missiles and torpedos.

This research aims to present a Bayesian projection model suitable for reliability growth planning of discrete-use systems suitable for use throughout all stages of system development. The model considers the three sources of uncertainty, including accounting for variability in the MS, vagaries in the mean FEF, and ambiguity in the initial system reliability. Traditional discrete-use models for reliability growth utilize test data from individual test events at the current stage of development. Little consideration is typically given to examining and using data from the previous bench, subsystem, or user trial events to periodically update the reliability growth model.

Specifically, the research aims to encompass the following points:

1. Establish and present a Bayesian modeling methodology for reliability growth projection applicable to discrete-use systems that considers uncertainty in the planned initial system reliability, the proposed MS, and the FEF.

2. Propose and test, utilizing representative real-world information, a methodology for combining multi-source data using the proposed projection model.

3. Extend the proposed Bayesian projection model to enable life cycle reliability planning, modeling, and projection throughout developmental, acceptance, and operational testing stages and in-service life.

4. Develop suitable open-source code based on the Python language, which permits reliability program development via simulation failure and corrective action approach.

1.3 *Research Contributions*

The proposed research aims to contribute to the literature by developing several novel approaches listed below.

1. A new Bayesian reliability projection model is proposed that considers the uncertainty surrounding discrete-use systems under arbitrary corrective action regimes to address failure modes. This differs from current models that fail to address the arbitrary nature of corrective action application strategies observed in real-world test situations. Additionally, the proposed strategy permits a probabilistic assessment of the test program, accounting for uncertainty in initial reliability and management variables.

2. An extension to the proposed Bayesian discrete-use projection model is developed by considering developmental, acceptance, and operational testing aspects through simulation of failures and corrective actions. This allows the formulation of a holistic reliability growth plan framework that encompasses the entire system lifecycle. The approach considers the posterior distribution

from each phase of developmental testing as the prior for the following growth test event. The same methodology is employed using the posterior from the final phase of reliability growth testing as the acceptance testing prior. It then follows that the acceptance testing posterior distribution forms the prior for subsequent operational testing. Importantly the approach is flexible enough to permit the combination of data from any test activity conducted in any order. The approach reduces unrealistic and unattainable reliability testing that may result from a purely statistical analysis. The proposed methodology also permits planning for combined developmental and operational test activities within a financially constrained context.

3. The research presents an approach for combining disparate data from various sources to establish prior distributions on system reliability.

4. The research develops and presents novel methods to assist reliability engineers in communicating developmental system reliability growth plans and risks to decision makers more effectively. The research takes essential facets of communication theory from marketing, management, business, and advertising and adapts them to the reliability engineering endeavor.

1.4 *Research Overview*

The following paragraphs briefly outline relevant chapters of the research dissertation.

Chapter 2 presents a brief overview of the literature survey conducted, including examining the historical and current state of the art in the methodologies developed for reliability growth planning and projection. The literature survey examines both continuous-use and discrete-use systems and considers potential modeling development and improvement opportunities.

Chapter 3 describes and outlines the proposed Bayesian model approach for discrete-use system reliability growth projection. The chapter centers on the Bayesian mathematical model development and includes a preliminary analysis of the proposed approach with noted advantages and disadvantages.

Annex A to Chapter 3 presents a detailed examination of selected failure and corrective action simulation results attained utilizing the methodology outlined in Chapter 3. The approach is compared to several other methods to determine suitability and robustness.

Chapter 4 presents a comparison of two practical empirical Bayes hyperparameter approaches useful for discrete-use system initial reliability estimation at the commencement of system-level developmental testing. The first was offered by Hall et al. [9], while the second was proposed and developed in Chapter 3.

Chapter 5 demonstrates and describes a hypothetical case study where the methods outlined in Chapter 3 and Chapter 4 are used to develop a complete reliability program for a representative discrete-use system. The program test phase projections developed from the proposed model are demonstrated as beneficial and improved compared to current methodologies.

Chapter 6 examines the issues surrounding the communication of reliability growth information, particularly the information flow requirements and methodologies of communication between reliability practitioners and decision makers. In particular, methods for effectively communicating developmental reliability growth plans and risks, such as those developed in Chapter 3, Chapter 4, and Chapter 5, are examined. Several new approaches are proposed that add to the available literature.

Chapter 7 completes the thesis work and includes conclusions, contributions, and extensions that may be considered for future exploration.

1.5 *Conclusions*

The conducted research offers insights into several important aspects of reliability growth management in terms of planning and projection, including:

1. the application of Bayesian modeling approaches to discrete-use system reliability program development,

2. the importance of combining multi-source data to enhance and improve reliability assessments through information use and reuse efficiencies,

3. methods useful in developing holistic approaches to extended reliability program development, and

4. the requirement to advance reliability growth communication.

Each of these aspects is important in evolving reliability growth methodologies and principles in a positive manner.

Chapter 2: Literature Review

2.1 *Introduction*

Many reliability growth modeling and assessment techniques have been developed in the history of reliability engineering. Typically, these models and techniques will fall into one of three categories:

- reliability growth planning models,
- reliability growth tracking models, or
- reliability growth projection models [10].

The purpose of the literature review is to provide a synopsis of the methodology and concepts that have been developed to assist in reliability growth planning and projection activities as they relate to the research.

For further detailed literature reviews that include reliability growth tracking models, reliability demonstration methods, system reliability assessments, reliability growth handbooks, and other relevant literature, the author directs the audience to Hall's 2008 [11] or Wayne's 2013 [12] doctoral dissertations. Both research publications feature detailed analyses of other reliability activities.

2.2 *Reliability Growth Planning*

A well-thought-out reliability growth plan can serve as a powerful management tool in scoping the required resources to enhance and demonstrate the system reliability requirement.

The following paragraphs describe a selection of the various reliability growth planning models, their development, and their utility.

2.2.1 Duane Model (1964)

The premise of the Duane Model [13] is the observation that the cumulative failure rate of the system with respect to cumulative test time is linear when examined on a logarithmic-logarithmic (or log-log) scale. The negative slope of the line indicates the rate of system reliability improvement and is known as the "growth rate." Alternatively, the Duane Model may be used to plot cumulative Mean Time Between Failure (MTBF) versus cumulative test time. In this case, the growth rate indicated by the slope of the line is positive. Several reliability growth models utilize the functional linear form of the Duane Model as a fundamental assumption. The model was developed initially to facilitate reliability growth monitoring and tracking in major subsystems for various aircraft; however, it has been crucial in developing later reliability growth planning and tracking models.

2.2.2 Selby and Miller's Reliability Planning and Management Model (1970)

The Selby and Miller Reliability Planning and Management (RPM) Model [14] was developed to plan and manage reliability programs for complex systems. The basic concept of the model follows the Duane Postulate. The Selby and Miller Model was the first known model to utilize the Duane Postulate for reliability growth planning.

The model follows the four Duane Postulate axioms:

- Reliability improvement of complex equipment follows a mathematically predictable pattern.
- Reliability improvement is approximately inversely proportional to the square root of cumulative operating (or test) time.
- For constant corrective action effort and implementation, reliability growth closely approximates a straight line on log-log scales.
- These relationships permit the use of a straightforward technique for monitoring progress toward a predetermined reliability goal.

2.2.3 Military Handbook 189 (MIL-HDBK-189) Model (1981)

The reliability growth planning model displayed with MIL-HDBK-189 [7] was first presented by Crow [15] in 1974. The model is based on the Power Law Non-Homogeneous Poisson Process (NHPP). It was the first stochastic use of the Duane Model and the second reliability growth planning model based on the Duane

Postulate. The model was designed to permit the development of a reliability growth plan over a multi-phase developmental test series. The model outlined an approach for reliability growth management during a system reliability growth program. The MIL-HDBK-189 Model was the first to outline an approach where the actual reliability growth could be compared to the planned growth.

The MIL-HDBK-189 Model develops an idealized Reliability Growth Curve (RGC) for a Test-Analyze-Fix-Test (TAFT) process. The TAFT process realizes reliability improvement when corrective actions are applied to failure modes observed during testing. The idealized growth curve is defined by five key parameters standard across all reliability growth planning models. These include the planned initial Mean Time Between Failure (MTBF), the length of the initial test phase, the goal MTBF to be achieved at the end of the test program, the reliability growth rate, and the total reliability growth test time. The model provides incremental reliability growth steps that show the individual test phase planned reliability targets based on the average value of the RGC over the Developmental Testing (DT) phases.

2.2.4 AMSAA System Level Planning Model (1992)

The United States Army Materiel Systems Analysis Activity (AMSAA) System Level Planning (SPLAN) Model is a variant of the MIL-HDBK-189 Model. It can be used to develop system reliability growth plans and associated idealized RGCs. Additionally, the model can also describe the required test duration to achieve a system reliability requirement as a point estimate. The model is often valuable for conducting sensitivity analysis; given any four of the five planning parameters within the MIL-HDBK-189 Model, SPLAN will determine the value of the remaining parameter. The initial MTBF, goal MTBF, growth rate, and length of the initial test phase are typically given and used to determine the total required test time.

2.2.5 AMSAA Subsystem Level Planning Model (1992)

The AMSAA Subsystem Level Planning (SSPLAN) Model [16] was designed to develop RGCs at the system or subsystem level where the MTBF requirement was to be demonstrated at the desired level of statistical confidence. The model may determine subsystem reliability performance requirements that align with the system reliability MTBF goal.

For the case where the system is modeled solely as one growth subsystem, SSPLAN is simply SPLAN. In this instance, system-level growth test planning parameters only are utilized. These inputs can be used to establish a target SPLAN idealized system-level RGC and associated test duration via the analytical formulas presented in the previous section. Alternately, the more general SSPLAN simulation procedure discussed in this section can be utilized.

The planning factors for SSPLAN are both system-level and subsystem level. The system-level planning factor is the system reliability requirement (or Technical Requirement, TR). The subsystem planning factors include the subsystem initial test time, t_i; the subsystem initial MTBF, $MTBF_i$; the Management Strategy, MS; and the probability of observing at least one B-mode failure for the subsystem, $Prob$. An array of planning strategies are available depending on available, known, or estimated subsystem information.

2.2.6 AMSAA Planning Model Based on Projection Methodology (2006)

The purpose of the AMSAA Planning Model Based on Projection Methodology (PM2) [17] is to produce a reliability growth plan to assist with reliability developmental testing management for complex systems that incorporate a developmental test schedule and a corrective action strategy. The critical difference between this model and other earlier reliability growth planning models is that it is independent of the NHPP assumption. PM2 does not utilize a growth parameter, and it considers planning parameters that can be directly influenced by reliability program management. These parameters include:

- the initial system MTBF, $MTBF_i$,
- Management Strategy, MS,
- the goal MTBF, $MTBF_g$,
- total developmental test time, T,
- average Fix Effectiveness Factor (FEF), μ_d,
- the number and placement of Corrective Action Periods (CAPs), and
- the average lag time associated with corrective actions.

Corrective action lag time can significantly impact the reliability improvement program, and PM2 is generally considered as the first planning model to include this impact.

PM2, like many of the other planning models, also utilizes an idealized growth curve and incremental steps that represent the goal reliability at each test phase. In PM2, however, the incremental steps also consider the lag time in applying corrective actions. This results in the individual steps falling below the idealized growth curve. These steps are based on the fact that most corrective actions for failure modes observed during test events are made during planned CAPs between individual test phases.

The PM2 methodology consists of two sub-models: PM2-Continuous, developed by Hall and Ellner in 2006 [17], and PM2-Discrete, developed by Hall in 2011 [9]. As their names highlight, one is suitable for planning continuous-use system reliability growth planning, while the other is suited for modeling discrete-use system reliability.

PM2 RGCs consists of two components: an idealized growth curve and reliability targets for each phase. Two types of failure modes are considered in the model parameters. A-mode failures are those failure modes that will not be addressed by planned corrective actions, whereas B-mode will. For both PM2 sub-models, the idealized curve may be interpreted as the expected system MTBF at test time $t \in [0, T]$ that would be achieved if all B-modes surfaced by t were addressed via corrective actions implemented with the planned average FEF where T represents the total test time. The idealized curve extends from the initial MTBF to the goal MTBF.

2.2.7 Crow Extended Model for Reliability Growth Planning (2010)

The Crow Extended Planning Model [18] is an improved version of the MIL-HDBK-189 Model. It utilizes many of the advancements of the PM2 methodology, and many of the inputs are the same. An additional input, known as the discovery Beta, is also required, which describes the discovery rate of correctible modes during testing. Due to the inherent connection with the NHPP associated with the MIL-HDBK-189 Model, a discovery Beta value of less than one indicates reliability growth is occurring.

2.2.8 Hall's Discrete Planning Model Based on Projection Methodology (2011)

Hall's Planning Model Based on Projection Methodology – Discrete (PM2-Discrete) [9] was developed as a discrete-use system analog to the original PM2 Model developed for continuous use. The model is based on the underlying reliability growth projection methodology to address the lack of projection models for discrete one-shot type systems. The model provides several uses for reliability program managers, including determining planned reliability achievement for available resources, acting as a baseline which realized reliability values could be measured against, and assessing the feasibility of the test program for achieving final reliability goals. The model also provides some valuable metrics associated with the reliability growth of the system.

PM2-Discrete utilizes similar planning parameters directly influenced by program management compared to PM2-Continuous, including the planned initial system reliability, the management strategy, the planned average fix effectiveness factor, the total developmental testing duration, and the average delay associated with corrective actions.

2.2.9 Wayne's Improved Planning Model Based on Projection Methodology – Continuous (2013)

The Wayne Improved PM2-Continuous Model further develops Hall's PM2-Continuous Model. In his 2013 dissertation [19], Wayne proposed an approach to explicitly combined developmental testing and operational testing to develop the MTBF goal.

Wayne further extended the approach in 2018 [20] by providing a Bayesian-based probabilistic treatment of various parameters. This is necessary to quantify the uncertainty present in the initial system reliability and management metrics. The technique also allows arbitrary time consideration for corrective actions. The approach allows for a reduction in the amount of demonstration testing necessary for a given level of uncertainty. It may also reduce high demonstration testing goals that typically result from a statistical OC Curve analysis.

Figure 1 displays an RGC developed utilizing Wayne's approach for a representative real-world continuous-use electromechanical system noting the development of uncertainty bounds across the various test phases.

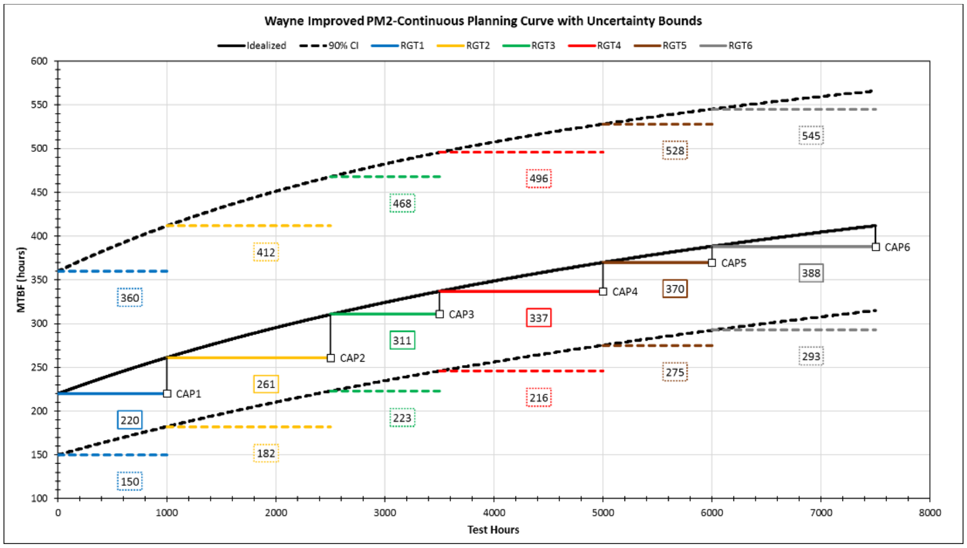


Figure 1. Example RGC developed using the Wayne continuous-use methodology.

Uncertainty may be present within the MS and FEF parameters. Wayne assumed that both variables could be modeled through a Beta distribution such that $MS \sim \text{Beta}(a_m, b_m)$ and $\mu_d \sim \text{Beta}(a_\mu, b_\mu)$ Treating these variables probabilistically resulted in the reliability mean and variance expressions of

$$E[\lambda_{DT}|t] = \frac{\left(\frac{b_m}{a_m + b_m}\right)\lambda_I t}{\left(\frac{1}{\beta} + t\right)} + \frac{\left(\frac{b_\mu}{a_\mu + b_\mu}\right)\left(\frac{a_m}{a_m + b_m}\right)\lambda_I t}{\left(\frac{1}{\beta} + t\right)} + \frac{\lambda_I}{1 + (\beta)t} \tag{1}$$

and

$$Var[\lambda_{DT}|t] = \frac{\left(\frac{b_m}{a_m + b_m}\right)\lambda_I t}{\left(\frac{1}{\beta} + t\right)^2} + \frac{\left(\frac{a_m}{a_m + b_m}\right)\left[\frac{b_\mu(b_\mu + 1)}{(a_\mu + b_\mu)(a_\mu + b_\mu + 1)}\right]\lambda_i t}{\left(\frac{1}{\beta} + t\right)^2} + \frac{\lambda_i}{\beta\left(\frac{1}{\beta} + t\right)^2} \tag{2}$$

where the β parameter is defined by Wayne as

$$\beta = \frac{\lambda_i - \lambda_g}{\lambda_g T_{DT} - \left[\left(\frac{b_m}{a_m + b_m}\right) + \left(\frac{b_\mu}{a_\mu + b_\mu}\right)\left(\frac{a_m}{a_m + b_m}\right)\right] \lambda_i T_{DT}} \quad (3)$$

2.3 *Reliability Growth Projection*

The reliability growth process applied to a complex system undergoing development involves surfacing failure modes, analyzing the modes and implementing corrective actions. In this manner, the system configuration is matured with respect to reliability. The rate of improvement in reliability is determined by:

- the ongoing rate at which new problem modes are being surfaced,
- the effectiveness and timeliness of corrective actions, and
- the set of failure modes that are addressed by corrective actions.

At the end of a test phase, program management usually desires an assessment of the system's reliability associated with the current configuration. Often, the amount of data generated from testing the current system configuration is severely limited. In such circumstances, if the failure data generated over several system configurations are consistent with a reliability growth model, we can pool the data over the tested configurations to estimate the parameters of the growth model. This, in turn, will yield a reliability tracking curve that gives estimates of the configuration reliabilities. The resulting assessment of the system's current reliability is called a demonstrated estimate since it is based solely on test data.

If the current configuration is the result of applying a group of fixes to the previous configuration, there could be a statistical lack of fit in tracking reliability growth between the previous and current configurations. In such a situation, it may not be valid to use a reliability growth tracking model to pool configuration data to assess the reliability of the current configuration. The option exists of estimating the current configuration reliability-based only on failure data generated for this configuration. However, such an estimate may be inadequate if little test time has been accumulated since the corrective actions were implemented. In this situation, program management may wish to use a reliability projection method. Such methods are typically based on assessing the effectiveness of corrective actions and failure data generated from the current and previous configurations.

A second situation in which a reliability projection is often utilized is when a group of corrective actions is scheduled for implementation at the end of the current test phase, before commencing a follow-on test phase. Program management often desires a projection of the reliability that will be achieved by implementing the delayed fixes. This type of projection can be based solely on the current test phase failure data and engineering assessments of the effectiveness of the planned corrective actions.

The following paragraphs describe the historical development of several reliability growth projection models.

2.3.1 Corcoran, Weingarten, and Zehna Model (1964)

Corcoran, Weingarten, and Zehna developed the first model for estimating reliability after corrective action [21]. The approach was developed considering estimating reliability in the final stage of development of an "expensive item." The reliability projection was suitable when corrective actions are undertaken at the end of a test consisting of N independent trials. The trial outcomes were assumed to follow a multinomial distribution with parameters N (total number of trials), q_0 (unknown success probability) and p_i (unknown failure probability for failure mode $i = 1, \cdots, k$). The assumption of a multinomial model implied that at most, one failure mode can occur per trial. An exact expression for the system reliability was presented, and comparisons of various estimators were provided.

2.3.2 AMSAA Crow Projection Model (1982)

The AMSAA Crow Projection Model used the NHPP interpretation of the Duane Postulate to describe the rate of occurrence of failure modes in the system [22]. The model intended to project the growth in reliability that would be seen at the beginning of the next testing phase following the implementation of planned corrective actions. The model assumed that all corrective actions were delayed until the end of the current test phase. The model is also one of the first to introduce the concept of the reliability growth potential, which was defined as the theoretical upper limit of reliability that could be achieved via the test-fix-test methodology. This concept remains an important factor that governs reliability growth programs in general. It is commonly considered and monitored for current reliability growth programs in the U.S. Department of Defense and the Australian Department of Defence. Two separate goodness-of-fit procedures were discussed, the Cramer Von-Mises and Chi-Squared tests; however, no interval procedures were described.

2.3.3 AMSAA Maturity Projection Model (1995)

The AMSAA Maturity Projection Model (AMPM) used a "doubly-stochastic" process to describe the underlying behavior of the system failure intensity [23] rather than a direct NHPP assumption. The model assumed that the system was comprised of a number of failure modes, with the collection of failure mode rates being realizations of a gamma distribution. The time between failures for each mode was then assumed to be exponential. AMPM was the first projection model to allow for arbitrary corrective actions, as the corrective actions could occur during the test or be delayed until after the test. As a result, it used only the first occurrences of each failure mode to develop failure intensity estimates. The AMPM was also the underlying methodology for developing the PM2 reliability growth planning model [24].

In addition to the system-level failure intensity, the model also provided estimates for the expected number of observed failure modes in later testing, the rate of occurrence of new failure modes, and the percent of the initial failure intensity on the modes that had been observed. Goodness of fit procedures were available using the expected number of failure modes; however, no confidence intervals have been developed.

2.3.4 Clark's Projection Model (1999)

The projection model proposed by Clark was developed due to the recognition that many programs failed to achieve significant reliability growth until late in the program nearer production [24]. Clark proposed that the reasoning for this occurrence was the lack of focus on reliability early in the development of a new system. The

Clark Model was an extension of the AMSAA Crow Projection Model [22] with two main differences. The first was that the original model was modified to allow for arbitrary corrective actions. The second was the addition of an inherent failure rate term that allows for decisions to be made regarding future reliability investment. Suppose the current reliability was too close to the maximum possible value. In that case, it might not have proven cost-effective to continue to invest in further reliability improvement through test-fix-test strategies.

2.3.5 AMSAA Maturity Projection Model – Stein (2004)

The AMSAA Maturity Projection Model - Stein (AMPM-Stein) [25] was developed as an extension to AMPM [23]. The extension limits one of the original assumptions of the model as corrective actions, in this case, must be delayed until after the test. The model used the same underlying theoretical structure as the original AMPM, but additional data were used to develop the model estimates. All of the data, both first and repeat occurrence times, were used to develop shrinkage estimates, or Stein estimates [26], to develop the model. A benefit of the approach was that the use of additional data increased the accuracy of the resulting estimates. The shrinkage estimation minimized the mean square error value, which provided an immediate connection to Bayesian modeling using squared error loss functions. As with the original model, goodness of fit procedures were available, but no confidence interval methods were developed.

2.3.6 Crow Extended Model (2004)

The Crow Extended Model was developed to model arbitrary corrective action strategies using the previously existing AMSAA Crow NHPP modeling framework [27]. The Crow Extended Model was a straightforward combination of the AMSAA Tracking Model [28] and the AMSAA Crow Projection Model [22]. Failure Modes were classified using the A-mode and B-mode distinction, where A-modes were not addressed via corrective action. The B-modes, those modes addressed via corrective actions, were further divided into BC-modes and BD-modes. Specifically, BC-modes have corrective actions implemented during the test phase, while BD-modes have corrective actions delayed until after the test phase is complete. The model used all A, BC, and BD-mode failures in the AMSAA Tracking Model to estimate the reliability growth that occurred during the test. The BD-mode failure intensity was then estimated using the maximum likelihood estimate, n_{BD}/T, for n_{BD} failures in test time T. Because the BD-mode corrective actions were delayed, their growth contribution during the test was subtracted from the Tracking Model estimate and replaced with a more appropriate estimator. The BD-mode failure intensity after corrective action was then estimated with the AMSAA Crow Projection Model. The overall result for the Crow Extended Model then subtracted the BD-mode maximum likelihood estimate n_{BD}/T from the AMSAA Tracking Model result and replaced it with the AMSAA Crow Projection Model result.

The model was shown via simulation study to provide extremely optimistic results when a large proportion of corrective actions were delayed [29]. There was also a logical discrepancy by treating the A and BD-modes together with the BC-modes with the AMSAA Tracking Model, which assumed that reliability growth was occurring simply because failure modes were being addressed during the test. The attempt to overcome the issue by subtracting out the BD-mode contribution resulted in a bias in the model that tends to provide systemically optimistic results.

2.3.7 Hall Discrete Projection Model (2008)

The discrete reliability growth projection model proposed by Hall [30] [31] was a discrete-use system counterpart to the AMPM-Stein Model [25]. The model used Stein-estimation procedures [26] to develop shrinkage estimates for the system's failure intensity of unobserved failure modes. All corrective actions were delayed until the end of the current test phase, and more than one failure mode can occur on a given trial during the test.

The discrete method proposed by Hall used a geometric likelihood for the first occurrence trial of an observed failure mode, and the mode probabilities of failure were assumed to be a realization from an underlying Beta distribution. Both methods of moments and maximum likelihood estimators were provided. Results were developed for systems with a known number of failure modes and those assumed to be complex with a large number of modes. Several associated management metrics were also presented, such as the expected number of new failure modes to be observed during additional testing, the rate of occurrence of new failure modes, and the reliability growth potential of the system. Model performance was also studied via

Monte Carlo simulation, and results indicated that performance was reasonable with only minor errors in the projection estimates.

2.3.8 Hall and Mosleh's Bayesian Methodology for Discrete Reliability Growth (2009)

The discrete reliability growth methodology presented by Hall and Mosleh [32] was developed as an additional estimation procedure to those first presented in [30] and [33]. Again, the approach used the underlying theoretical assumption of mode failure probabilities as realizations from an underlying Beta distribution. Additional assumptions include a Binomial distribution for observed failures during the test, with all corrective actions delayed until the end of the testing.

The Bayesian inference in the model was used only to estimate the parameters of the underlying Beta distribution. Squared error loss was used along with a constant prior, and numerical methods were utilized to evaluate the resulting posterior. Simulation methods were also employed to generate uncertainty distributions on each of the management metrics developed in [30] and [31].

2.3.9 Hall, Ellner, and Mosleh's Discrete Reliability Growth Projection Model (2010)

The model presented in this paper used the underlying assumption of mode failure probabilities as realizations from a Beta distribution [33]. The model differs from the previously presented discrete models somewhat, as it allows for arbitrary corrective actions to occur either during or directly after the test. Only failure mode first occurrence trials were used, along with the corresponding FEF for each failure mode. Because only the first occurrence of each failure mode was used, the geometric distribution was used to model the mode's first occurrences. Goodness-of-fit procedures were presented in order to validate the model assumptions.

Maximum likelihood estimates were developed for the Beta distribution parameters, with results given for a finite number of failure modes and a complex system consisting of a large number of failure modes. Several management metrics and model equations were also developed, such as the reliability growth potential, the expected number of new failure modes, and the fraction of the initial failure probability surfaced during the testing.

2.4 *Conclusions*

This chapter reviewed a number of reliability growth planning and projection models found in the literature to define the current state-of-the-art. Many modeling approaches are documented in the literature for both continuous-use and discrete-use systems, with several classical and Bayesian approaches available.

From reviewing the literature, it is evident that there is a general lack of reliability growth approaches that consider data from throughout the developmental program of the system. This is particularly true for reliability growth projection models. The models apply to single test phases only, with no way of updating results from test phase to test phase. The models involving arbitrary corrective action strategies can also be improved, as the current state-of-art involves using only the first observed time or trial of occurrence for a given failure mode.

There are also limited approaches for combining data from different types of testing within a reliability growth program. Current approaches for data combination from different test modalities involve various types of testing in combination with reliability development testing. However, there are limited options that consider reliability growth testing combined with operational or reliability demonstration testing. The failure mode-based options currently available in the literature also consider only finite numbers of known failure modes in the system, with no allowance for unobserved failure modes. These approaches are also developed specifically for time-to-failure distributions, with no extensions to consider reliability for complex repairable systems.

Finally, there is a decided disconnect between the current reliability growth approaches in the literature and the reliability assessment methods involving reliability engineering efforts. The use of component and subsystem data for system reliability assessment occurs in various papers. However, none of these discuss an approach that connects the results to reliability growth modeling approaches. The use of physics-based results within these approaches is also limited.

Chapter 3: A Bayesian Approach to Discrete-Use System Reliability Growth Planning and Projection Under Arbitrary Corrective Actions[1]

3.1 *Abstract*

The term discrete-use describes systems whose usage is measured discrete demands such as missile systems and torpedoes. As discrete-use systems developmental programs address increasingly complex platforms, the risk, and complexity associated with evolving a sound reliability growth plan also increase. Good reliability growth management is foundational in ensuring reliability achievement as a function of available time and resource constraints. However, current discrete-use system reliability models typically do not consider several uncertainties associated with the reliability growth plan. These include management planning parameters, the variability in corrective actions designed to address observed failure modes, and the underlying mode hazard rate. Consequently, this increases the risk associated with the reliability growth program and the overall developmental product schedule. This chapter presents an approach to reliability growth planning and projection that considers these uncertainties resulting in a practical and real-world representative

---

[1] The content of this chapter was presented at the Australian Integrated Project Engineering Congress (IPEC) May 26-28, 2021. Additionally, an adapted version of this chapter was submitted in the American Society of Mechanical Engineers (ASME) Safety Engineering, Risk and Reliability Analysis Division (SERAD) 2021 Student Paper on Safety Innovation Contest and was awarded an Honorable Mention. A full-text adapted version of this chapter has been submitted for publishing consideration to the Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, a SAGE Publishing publication.

model. The proposed method is a natural extension of current reliability growth planning and projection models used under developmental test conditions. The approach models the uncertainty in the underlying failure mode probability, the failure mode corrective action effectiveness, and the management strategy for dealing with observed failure modes within a Bayesian framework. Uncertainty quantification is essential in ensuring decision makers are aware of the range of potential consequences and the breadth of risk that various management strategies and corrective action schemes entail.

3.2 *Introduction*

Reliability growth is defined as the improvement in a reliability metric of a component, subsystem, or system over a period due to changes in the product's design, the adopted manufacturing processes, or both. The objective of reliability growth testing is to prove increases in a system's reliability to a particular goal or requirement in a phased manner by discovering failure modes and implementing corrective actions [34]. All planning and projection models have limitations and often fail to represent or oversimplify the true nature of testing and growth activities. Early models typically assumed that the observed failure modes and their associated corrective actions were incorporated during the test and at the specific failure time. These methods produce an "idealized growth curve" [7]. In practice, corrective action implementation may be limited during testing due to schedule, availability of engineering resources, the appropriateness of test data, or other constraints. Corrective actions that require more investigation may be delayed until after the completion of the current test phase. Some failure modes may also not be addressed, depending on the management strategy adopted at any singular point in time.

This chapter presents a framework for discrete-use system reliability growth planning and projection under real-world mixed arbitrary corrective actions. It considers additional elements of uncertainty previous models have not included: unpredictability in the initial system reliability upon entry to the test program, variability in the management strategy adopted in dealing with observed failure modes, and inconsistency in corrective action effectiveness. The model presented is a Bayesian approach which aligns with elements of Hall's Discrete Projection Model [30] [31] and also Hall, Ellner, and Mosleh's Discrete Reliability Growth Projection Model [33]. Extension of the basic model also provides a useful alternative to a contemporary discrete-use system growth planning model: Planning Model Based on Projection Methodology – Discrete (PM2-D) [7]. The proposed approach combines the benefits realized from Bayesian approaches and simulation techniques to deliver realistic and detailed reliability growth plans.

3.3 *Model Assumptions*

The modeling approach adopted follows similar methods described by others [35] [36] in that system reliability is considered a product of reliabilities from independent failure modes. The effectiveness of corrective actions can be mathematically quantified when the system is modeled and enables modal fix effectiveness factors (FEF) to address observed failure modes. FEF values represent the fraction of a failure mode's rate of occurrence mitigated by corrective action. An FEF of one represents an idealized "perfect" corrective action resulting in complete mitigation of the observed failure mode. Conversely, an FEF of zero indicates an entirely ineffective corrective action that does not change the mode failure probability of occurrence [37].

In developing the discrete-use system model, the following assumptions were made:

1. The discrete-use system consists of many serial failure modes, with the occurrence of any failure mode resulting in system failure.

2. Failure modes are independent of each other.

3. The mode failure probability remains constant before and after the implementation of corrective actions.

4. The implemented corrective action has an associated FEF resulting in a fraction mitigation of the underlying constant mode failure probability.

5. The applied corrective actions do not introduce new failure modes.

3.4 *Single Failure Mode Posterior Inference*

To develop the model of system reliability, we consider the posterior distribution for a single failure mode. For a given test duration of T total demands with arbitrary corrective actions, assume for the i^{th} failure mode, there are n_i total failures on trials $t = (t_{i,1}, \dots, t_{i,ni})$ with corrective action implemented on trial v_i and an FEF of d_i. Additionally, there are $n_{i,1}$ mode failures before corrective action implementation and $n_{i,2} = n_i - n_{i,1}$ mode failures after corrective action.

Figure 2 displays the single failure mode and corrective action concept in number line form.

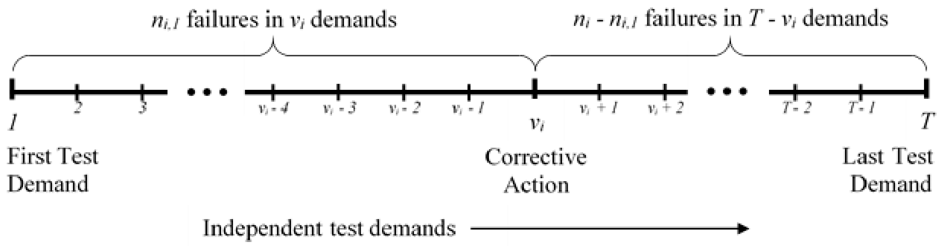


Figure 2. Single failure mode and corrective action concept.

Let p_i denote the mode failure probability, then for a series of independent Bernoulli (or Binomial) trials, the probability mass function (PMF) for the observation of the single failure mode before corrective action is

$$f(n_{i,1}, v_i, p_i) = \binom{v_i}{n_{i,1}} p_i^{n_{i,1}} (1 - p_i)^{v_i - n_{i,1}} \tag{4}$$

with the $p_i^{n_{i,1}}$ term representing observed failures and $(1 - p_i)^{v_i - n_{i,1}}$ representing test demand successes with $\binom{v_i}{n_{i,1}}$ different ways of distributing $n_{i,1}$ failures in v_i trials.

The PMF, after the implementation of the corrective action, is then

$$f(T, n_i, n_{i,1}, v_i, p_i) = \binom{T - v_i}{n_i - n_{i,1}} (1 - p_i)^{(1-d_i)[T - v_i - (n_i - n_{i,1})]} [1 - (1 - p_i)^{1-d_i}]^{n_i - n_{i,1}} \tag{5}$$

with the $(1 - p_i)^{(1-d_i)[T - v_i - (n_i - n_{i,1})]}$ term representing additional test demand successes and the $[1 - (1 - p_i)^{1-d_i}]^{n_i - n_{i,1}}$ term further additional observed instances of the same failure mode after corrective action. Correspondingly, the $\binom{T - v_i}{n_i - n_{i,1}}$ term represents the different ways of distributing $n_i - n_{i,1}$ failures in $T - v_i$ trials.

With the assumptions of failure mode independence and a constant mode failure probability, these aspects result in a failure mode's likelihood function shown below, such that

$$\ell(t|p_i, v_i, d_i, T) \propto p_i^{n_{i,1}}(1 - p_i)^{v_i - n_{i,1}}(1 - p_i)^{(1-d_i)[T - v_i - (n_i - n_{i,1})]}[1 - (1 - p_i)^{1-d_i}]^{n_i - n_{i,1}} \qquad (6)$$

where

- $p_i^{n_{i,1}}$ represents trial failures before corrective action
- $(1 - p_i)^{v_i - n_{i,1}}$ represents trial successes before corrective action
- $(1 - p_i)^{(1-d_i)[T - v_i - (n_i - n_{i,1})]}$ represents trial successes after corrective action
- $[1 - (1 - p_i)^{1-d_i}]^{n_i - n_{i,1}}$ represents trial failures after corrective action.

A Beta(a,b) distribution is selected as a useful representation of the prior distribution on the mode failure probability such that

$$p(p_i) = \frac{\Gamma[a + b]}{\Gamma[a]\Gamma[b]}\, {p_i}^{a-1}(1 - p_i)^{b-1} \qquad (7)$$

The Beta distribution is suitable as it is the conjugate prior distribution for the Bernoulli, Binomial, negative Binomial and geometric distributions and can be used to appropriately model the behavior of random variables limited to the finite interval [0, 1].

The mode failure posterior distribution is then

$$p(p_i|t) = \frac{p(p_i)\ell(t|p_i, v_i, d_i, T)}{\int_0^1 p(p_i)\,\ell(t|p_i, v_i, d_i, T)dp_i} \tag{8}$$

Solving the posterior requires substituting $n_{i,2} = n_i - n_{i,1}$ and employing a Binomial expansion such that

$$[1-(1-p_i)^{1-d_i}]^{n_{i,2}} = \sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j (1-p_i)^{(1-d_i)j} \tag{9}$$

Equation (9) results in a failure mode posterior distribution of

$$p(p_i|t) = \frac{\sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j {p_i}^{a+n_{i,1}-1} (1-p_i)^{b+v_i-n_{i,1}+(1-d_i)(T-v_i-n_{i,2}+j)-1}}{\sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j \frac{\Gamma[a+n_{i,1}]\Gamma[b+v_i-n_{i,1}+(1-d_i)(T-v_i-n_{i,2}+j)]}{\Gamma[a+b+v_i+(1-d_i)(T-v_i-n_{i,2}+j)]}} \tag{10}$$

If we assume that all corrective actions are delayed until the end of the test, then $n_{i,2} = 0$ as there can be no observed single mode failures after corrective action and $v_i = T$. If this is the case, then the failure mode posterior distribution in Equation (10) simplifies to the standard Beta-Binomial conjugate relationship such that

$$p(p_i|t) = \frac{{p_i}^{a+n_{i,1}-1}(1-p_i)^{b+T-n_{i,1}-1}}{\frac{\Gamma\left[a+n_{i,1}\right]\Gamma\left[b+T-n_{i,1}\right]}{\Gamma[a+b+T]}}$$

$$= \frac{\Gamma[a+b+T]}{\Gamma\left[a+n_{i,1}\right]\Gamma\left[b+T-n_{i,1}\right]}{p_i}^{a+n_{i,1}-1}(1-p_i)^{b+T-n_{i,1}-1} \tag{11}$$

Similarly, if no failures for the single mode are observed during the test, then Equation (10) simplifies to

$$p(p_i|t) = \frac{{p_i}^{a-1}(1-p_i)^{b+T-1}}{\frac{\Gamma[a]\Gamma[b+T]}{\Gamma[a+b+T]}}$$

$$= \frac{\Gamma[a+b+T]}{\Gamma[a]\Gamma[b+T]}{p_i}^{a-1}(1-p_i)^{b+T-1} \tag{12}$$

However, our interest remains in developing the posterior reliability for the failure mode. Using Equation (10), the posterior mean reliability is found as the expectation of $1-p_i|t$

$$E(1-p_i|t)$$
$$= \frac{\sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j \frac{\Gamma\left[b+v_i-n_{i,1}+(1-d_i)\left(T-v_i-n_{i,2}+j\right)+1\right]}{\Gamma\left[a+b+v_i+(1-d_i)\left(T-v_i-n_{i,2}+j\right)+1\right]}}{\sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j \frac{\Gamma\left[b+v_i-n_{i,1}+(1-d_i)\left(T-v_i-n_{i,2}+j\right)\right]}{\Gamma\left[a+b+v_i+(1-d_i)\left(T-v_i-n_{i,2}+j\right)\right]}} \tag{13}$$

The mean posterior failure mode reliability after corrective action implementation then becomes

$$E\left[(1-p_i)^{(1-d_i)}|t\right]$$

$$=\frac{\sum_{j=0}^{n_{i,2}}\binom{n_{i,2}}{j}(-1)^j\frac{\Gamma\left[b+v_i-n_{i,1}+(1-d_i)\left(T-v_i-n_{i,2}+j+1\right)\right]}{\Gamma\left[a+b+v_i+(1-d_i)\left(T-v_i-n_{i,2}+j+1\right)\right]}}{\sum_{j=0}^{n_{i,2}}\binom{n_{i,2}}{j}(-1)^j\frac{\Gamma\left[b+v_i-n_{i,1}+(1-d_i)\left(T-v_i-n_{i,2}+j\right)\right]}{\Gamma\left[a+b+v_i+(1-d_i)\left(T-v_i-n_{i,2}+j\right)\right]}} \quad (14)$$

3.5 <u>Complex System Posterior Inference</u>

A single failure mode is used to develop a similar system-level model for a complex system involving many failure modes. If we assume the system has k serial failure modes per our first assumption, then

$$E(R|t)=E\left[\prod_{i=1}^{k}(1-p_i)^{(1-d_i)}\,|t\right]$$

$$=\prod_{i=1}^{m}\frac{\sum_{j=0}^{n_{i,2}}\binom{n_{i,2}}{j}(-1)^j\frac{\Gamma\left[b+v_i-n_{i,1}+(1-d_i)\left(T-v_i-n_{i,2}+j+1\right)\right]}{\Gamma\left[a+b+v_i+(1-d_i)\left(T-v_i-n_{i,2}+j+1\right)\right]}}{\sum_{j=0}^{n_{i,2}}\binom{n_{i,2}}{j}(-1)^j\frac{\Gamma\left[b+v_i-n_{i,1}+(1-d_i)\left(T-v_i-n_{i,2}+j\right)\right]}{\Gamma\left[a+b+v_i+(1-d_i)\left(T-v_i-n_{i,2}+j\right)\right]}}$$

$$\times\left[1-\frac{a}{a+b+T}\right]^{k-m} \quad (15)$$

where m is the number of observed failure modes, and k is the total number of failure modes in the system. Equation (15) represents the product of mean failure mode reliabilities for the entire system. The equation complex right-hand side first term signifies the system observed failure modes derived as the product of many individual modes as detailed in Equation (14). The equation right-hand side second term represents the unobserved failure modes and is the mean of the posterior in Equation (12) when modes remain unobserved. Typically, we do not know the system's total failure mode number when considering a real-world scenario. To develop an estimate that does not include the total number of failure modes, we take the limit of Equation (15) as k becomes large. Before taking the limit, we reparameterize Equation (15) using the prior mean reliability of the system and an additional parameter. First, let

$$\tilde{n} = a + b \tag{16}$$

where $\tilde{n}$ is an additional parameter of the Beta(a,b) distribution mode failure probability. Then let the mean prior system-level reliability be denoted by

$$R_I = \prod_{i=1}^{k}\left(1 - \frac{a}{a+b}\right) = \left(1 - \frac{a}{a+b}\right)^k \tag{17}$$

which is the product of the prior mean failure mode reliabilities. The Beta distribution a parameter may then be written using Equation (16) and Equation (17) as

$$a = \tilde{n}\left(1 - {R_I}^{1/k}\right) \tag{18}$$

Examination of Equation (18) reveals that as k becomes large, $a \to 0$. This also implies that $b \to \tilde{n}$ as k becomes large.

Reparameterising Equation (15) through the substitution of Equation (18) now gives

$$E(R|t) = E\left[\prod_{i=1}^{k}(1-p_i)^{(1-d_i)}\,|t\right]$$

$$= \prod_{i=1}^{m} \frac{\sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j \frac{\Gamma\left[\tilde{n} - a + v_i - n_{i,1} + (1-d_i)\left(T - v_i - n_{i,2} + j + 1\right)\right]}{\Gamma\left[\tilde{n} + v_i + (1-d_i)\left(T - v_i - n_{i,2} + j + 1\right)\right]}}{\sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j \frac{\Gamma\left[\tilde{n} - a + v_i - n_{i,1} + (1-d_i)\left(T - v_i - n_{i,2} + j\right)\right]}{\Gamma\left[\tilde{n} + v_i + (1-d_i)\left(T - v_i - n_{i,2} + j\right)\right]}}$$

$$\times \left[1 - \frac{a}{\tilde{n} + T}\right]^{k-m} \tag{19}$$

Taking the limit of the right-hand side second term representing the unobserved failure modes in Equation (19) with respect to k gives

$$\lim_{k\to\infty} \left[1 - \frac{a}{\tilde{n} + T}\right]^{k-m} = \exp\left(\frac{\tilde{n}}{\tilde{n} + T} \log R_I\right) = R_I^{\frac{\tilde{n}}{\tilde{n}+T}} \tag{20}$$

The limit of the mean reliability of the complex system from Equation (19) is then

$$\lim_{k\to\infty} E(R|t)$$

$$= \lim_{k\to\infty} \prod_{i=1}^{m} \frac{\sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j \frac{\Gamma\left[\tilde{n} - a + v_i - n_{i,1} + (1-d_i)\left(T - v_i - n_{i,2} + j + 1\right)\right]}{\Gamma\left[\tilde{n} + v_i + (1-d_i)\left(T - v_i - n_{i,2} + j + 1\right)\right]}}{\sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j \frac{\Gamma\left[\tilde{n} - a + v_i - n_{i,1} + (1-d_i)\left(T - v_i - n_{i,2} + j\right)\right]}{\Gamma\left[\tilde{n} + v_i + (1-d_i)\left(T - v_i - n_{i,2} + j\right)\right]}}$$

$$\times \left[1 - \frac{a}{\tilde{n} + T}\right]^{k-m}$$

$$= \prod_{i=1}^{m} \frac{\sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j \frac{\Gamma\left[\tilde{n} - a + v_i - n_{i,1} + (1-d_i)\left(T - v_i - n_{i,2} + j + 1\right)\right]}{\Gamma\left[\tilde{n} + v_i + (1-d_i)\left(T - v_i - n_{i,2} + j + 1\right)\right]}}{\sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j \frac{\Gamma\left[\tilde{n} - a + v_i - n_{i,1} + (1-d_i)\left(T - v_i - n_{i,2} + j\right)\right]}{\Gamma\left[\tilde{n} + v_i + (1-d_i)\left(T - v_i - n_{i,2} + j\right)\right]}} \times R_I^{\frac{\tilde{n}}{\tilde{n}+T}} \qquad (21)$$

We must calculate the second moment for the posterior mean in Equation (14) to calculate the posterior variance. The second moment of the mean posterior is

$$E\left[\left\{(1-p_i)^{(1-d_i)}\right\}^2 | t\right]$$

$$= \frac{\sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j \frac{\Gamma\left[b + v_i - n_{i,1} + (1-d_i)\left(T - v_i - n_{i,2} + j + 2\right)\right]}{\Gamma\left[a + b + v_i + (1-d_i)\left(T - v_i - n_{i,2} + j + 2\right)\right]}}{\sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j \frac{\Gamma\left[b + v_i - n_{i,1} + (1-d_i)\left(T - v_i - n_{i,2} + j\right)\right]}{\Gamma\left[a + b + v_i + (1-d_i)\left(T - v_i - n_{i,2} + j\right)\right]}} \qquad (22)$$

The second moment of the posterior displayed in Equation (22) can now be further developed similarly to the posterior mean such that

$$E(R^2|t) = \prod_{i=1}^{m} \frac{\sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j \frac{\Gamma\left[\tilde{n} - a + v_i - n_{i,1} + (1 - d_i)\left(T - v_i - n_{i,2} + j + 2\right)\right]}{\Gamma\left[\tilde{n} + v_i + (1 - d_i)\left(T - v_i - n_{i,2} + j + 2\right)\right]}}{\sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j \frac{\Gamma\left[\tilde{n} - a + v_i - n_{i,1} + (1 - d_i)\left(T - v_i - n_{i,2} + j\right)\right]}{\Gamma\left[\tilde{n} + v_i + (1 - d_i)\left(T - v_i - n_{i,2} + j\right)\right]}}$$

$$\times \left[\frac{(b + T + 1)(b + T)}{(a + b + T + 1)(a + b + T)}\right]^{k-m} \tag{23}$$

Again, the complex right-hand side equation first term of Equation (23) denotes the system observed failure modes second moment derived as the product of many individual modes as detailed in Equation (22). Similarly, the right-hand side equation second term represents the unobserved failure modes second moment.

Using the same methodology demonstrated in Equation (16) to Equation (18), the complex system reliability contribution from unobserved failure modes can be expressed as

$$\left[\frac{(b + T + 1)(b + T)}{(a + b + T + 1)(a + b + T)}\right]^{k-m} = \left[\left(1 - \frac{a}{\tilde{n} + T + 1}\right)\left(1 - \frac{a}{\tilde{n} + T}\right)\right]^{k-m} \tag{24}$$

Taking the limit of Equation (24) with respect to k yields

$$\lim_{k \to \infty} \left[\left(1 - \frac{a}{\tilde{n} + T + 1}\right)\left(1 - \frac{a}{\tilde{n} + T}\right)\right]^{k-m} = R_I^{\frac{\tilde{n}}{\tilde{n}+T+1} + \frac{\tilde{n}}{\tilde{n}+T}} \tag{25}$$

Taking the limit of Equation (23) with respect to k and substituting the resultant from Equation (25) gives

$$\lim_{k\to\infty} E(R^2|t) = \prod_{i=1}^{m} \frac{\sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j \frac{\Gamma\left[\tilde{n} + v_i - n_{i,1} + (1-d_i)\left(T - v_i - n_{i,2} + j + 2\right)\right]}{\Gamma\left[\tilde{n} + v_i + (1-d_i)\left(T - v_i - n_{i,2} + j + 2\right)\right]}}{\sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j \frac{\Gamma\left[\tilde{n} + v_i - n_{i,1} + (1-d_i)\left(T - v_i - n_{i,2} + j\right)\right]}{\Gamma\left[\tilde{n} + v_i + (1-d_i)\left(T - v_i - n_{i,2} + j\right)\right]}}$$
$$\times R_I^{\frac{\tilde{n}}{\tilde{n}+T+1} + \frac{\tilde{n}}{\tilde{n}+T}} \tag{26}$$

Denoting the second moment in Equation (26) as μ_2 and the mean in Equation (21) as μ, the posterior variance for the complex system reliability is then identified as

$$Var[R|t] = \sigma^2 = \mu_2 - \mu^2 \tag{27}$$

Popular methods to represent the exact product of independent Beta distributions include the Meijer G or Fox's H functions [38] [39]; however, both have drawbacks regarding precision and computational resource usage in all commonly available software. Although the product of individual random Beta distribution variables does not typically follow an exact Beta distribution, we have found that the Beta distribution still forms a suitable approximation of the posterior for the system-level reliability without the computational resource concerns. A method of moments approach can be used to determine the parameters of the approximate distribution resulting in a set of simultaneous equations such that

$$\mu = \frac{a_1}{a_1 + b_1} \tag{28}$$

and

$$\sigma^2 = \frac{a_1 b_1}{(a_1 + b_1)^2 (a_1 + b_1 + 1)} \tag{29}$$

where a_1 and b_1 are the parameters for the Beta distribution approximation.

3.6 *Selecting Appropriate Reliability Growth Management Planning Parameters*

In undertaking sound reliability growth planning, two essential parameters must be considered that significantly influence the outcomes of reliability growth model simulations and real-world testing. The two planning parameters in question are the FEF and the Management Strategy (MS).

The adopted MS has a significant influence on the system reliability growth potential and the overall shape of the planned growth curve. Observation of a new failure mode during the growth program prompts a management decision to either do nothing and ignore the failure mode (known as Type A modes) or address the failure mode via corrective action (Type B modes). The MS parameter is the fraction of the initial system failure intensity planned to be in the Type B group. The MS parameter "…is a measure of how aggressive corrective actions are incorporated into the design" [40].

Many techniques for reliability growth planning treat the two key management planning parameters as a deterministic mean value. The approach presented here considers the uncertainty present in the parameter estimations through their treatment as continuous random variables. In both FEF and MS cases, a Beta(a,b) distribution is chosen to represent the respective parameters which are limited to the finite interval [0,1].

3.6.1 Principle of Maximum Entropy

The Principle of Maximum Entropy considers that when estimating the probability distribution, one should select that distribution that leaves the most significant remaining uncertainty consistent with the applicable constraints [41] [42]. Modeling the management planning parameters using the Principle of Maximum Entropy ensures approximate Beta(a,b) distributions representative of the expected FEF and MS encompassing the greatest uncertainty. Typically, a mean value for the parameter in question is assumed by examining historical data relating to similar systems, human factors and assessing management risk tolerance thresholds or analyzing the likely failure modes in the design. The mean is generally used for planning purposes, but other moments, such as variance or kurtosis, could also be used if available. This would change the form of the resulting distribution, which would either require a different method for approximating the Beta distribution or a new derivation using the actual distribution.

Maximizing the entropy subject to the assumed mean value within the range [0,1] results in the prior distribution being a truncated exponential distribution given by

$$p(x) = \frac{\mu \exp(-\mu x)}{1 - \exp(-\mu)} \tag{30}$$

where x is the mean value of the parameter of interest (either the FEF or MS), and μ is the solution of

$$\frac{1}{\mu} - \frac{\exp(-\mu)}{1 - \exp(-\mu)} = \varepsilon \tag{31}$$

for the mean parameter of interest value ε.

Examination of Equation (31) reveals that a discontinuity exists when the mean value equals 0.5. In reality, the mean parameter value is unlikely to be known with a high degree of precision. Consequently, altering the mean parameter value slightly (to 0.49 or 0.51) permits a positive solution. Assessment of results has demonstrated that the solution is insensitive to this minor adjustment.

A Beta(a,b) distribution can approximate the truncated exponential distribution in Equation (30) with reasonable accuracy. The Beta distribution parameters can be found by equating the mean and second moment about the origin of the two distributions resulting in the system of equations given by

$$\frac{a}{a+b} = \frac{1}{\mu} - \frac{\exp(-\mu)}{1 - \exp(-\mu)} = \varepsilon \tag{32}$$

and

$$\left(\frac{a}{a+b}\right)\left(\frac{a+1}{a+b+1}\right) = \frac{-\exp(-\mu) + 2\left[-\frac{1}{\mu}\exp(-\mu) - \frac{1}{\mu^2}\exp(-\mu) + \frac{1}{\mu^2}\right]}{1 - \exp(-\mu)} \tag{33}$$

3.7 *Prior Distribution Empirical Bayes Estimators*

Procedures exist to parameterize the prior Beta distribution on the reliability by exploiting historical data or eliciting expert opinion [43] when the proposed approach is used for reliability planning or actual failure data when the approach is used for reliability growth projection to make an estimate. Alternatively, empirical Bayes estimates can be developed for the proposed approach.

Note that the mean reliability in Equation (21) and the variance in Equation (27) are expressed in terms of prior system-level mean reliability R_I and the $\tilde{n}$ parameter. The empirical Bayes estimators for these parameters are developed by examining the likelihood in Equation (6) and the marginal likelihood when all failure modes are considered. The marginal likelihood for a single failure mode is the denominator of the posterior distribution in Equation (10) written as

$$p(n_i) = \frac{\Gamma[a+b]}{\Gamma[a]\Gamma[b]} \sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j \frac{\Gamma\left[a+n_{i,1}\right]\Gamma\left[b+v_i-n_{i,1}+\tau_{i,j}\right]}{\Gamma\left[a+b+v_i+\tau_{i,j}\right]} \tag{34}$$

where $\tau_{i,j} = (1-d_i)\left(T - v_i - n_{i,2} + j\right)$ and n_i remains the total number of observed failures during the test phase for the i^{th} failure mode.

From Equation (34), the total likelihood when considered over k modes in the complex system is

$$\mathcal{L}(n) = \prod_{i=1}^{k} \frac{\Gamma[a+b]}{\Gamma[a]\Gamma[b]} \sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j \frac{\Gamma\left[a+n_{i,1}\right]\Gamma\left[b+v_i-n_{i,1}+\tau_{i,j}\right]}{\Gamma\left[a+b+v_i+\tau_{i,j}\right]} \tag{35}$$

and the subsequent log-likelihood is

$$l(n) = \prod_{i=1}^{k} \Bigg\{ \log\Gamma[a+b] - \log\Gamma[a] - \log\Gamma[b]$$

$$+ \log\left[\sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j \frac{\Gamma\left[a+n_{i,1}\right]\Gamma\left[b+v_i-n_{i,1}+\tau_{i,j}\right]}{\Gamma\left[a+b+v_i+\tau_{i,j}\right]}\right]\Bigg\} \tag{36}$$

If we again assume that m failure modes are observed during the test phase, then the log-likelihood in Equation (36) may be represented as

$$l(n) = \log\left[\binom{k}{m} \prod_{i=1}^{m} \frac{\Gamma[a+b]}{\Gamma[a]\Gamma[b]} \sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j \frac{\Gamma\left[a+n_{i,1}\right]\Gamma\left[b+v_i-n_{i,1}+\tau_{i,j}\right]}{\Gamma\left[a+b+v_i+\tau_{i,j}\right]}\right]$$

$$+(k-m)\log\frac{\Gamma[a+b]\Gamma[b+T]}{\Gamma[a+b+T]\Gamma[b]} \tag{37}$$

Equation (37) is the sum of log-likelihood terms for m observed failure modes and $k-m$ unobserved failure modes. The constant $\binom{k}{m}$ represents the possible ways of observing m modes from the total population of k failure modes. Reparameterising Equation (37) in terms of R_I and the $\tilde{n}$ parameter gives

$$l(n) = \log\left[\binom{k}{m}\prod_{i=1}^{m}\frac{\Gamma[\tilde{n}]\Gamma\left[\tilde{n}\left(1-{R_I}^{1/k}\right)+n_{i,1}\right]}{\Gamma\left[\tilde{n}\left(1-{R_I}^{1/k}\right)\right]\Gamma\left[\tilde{n}{R_I}^{1/k}\right]}\sum_{j=0}^{n_{i,2}}\binom{n_{i,2}}{j}(-1)^j\frac{\Gamma\left[\tilde{n}{R_I}^{1/k}+v_i-n_{i,1}+\tau_{i,j}\right]}{\Gamma\left[\tilde{n}+v_i+\tau_{i,j}\right]}\right]$$

$$+(k-m)\log\frac{\Gamma[\tilde{n}]\Gamma\left[\tilde{n}{R_I}^{1/k}+T\right]}{\Gamma[\tilde{n}+T]\Gamma\left[\tilde{n}{R_I}^{1/k}\right]} \tag{38}$$

Taking the limit of Equation (38) as k becomes large yields

$$l_\infty(n) = \lim_{k\to\infty} l(n)$$

$$= \sum_{i=1}^{m}\log\left(\log {R_I}^{\tilde{n}}\prod_{\mathrm{q}=1}^{n_{i,1}-1}\left(n_{i,1}-q\right)\sum_{j=0}^{n_{i,2}}\sum_{j=0}^{n_{i,2}}\binom{n_{i,2}}{j}(-1)^j\frac{\Gamma\left[\tilde{n}+v_i-n_{i,1}+\tau_{i,j}\right]}{\Gamma\left[\tilde{n}+v_i+\tau_{i,j}\right]}\right)$$

$$+\log {R_I}^{\tilde{n}}[\Psi(\tilde{n}+T)-\Psi(\tilde{n})] \tag{39}$$

where Ψ is the Digamma function. Taking the derivative with respect to R_I and $\tilde{n}$ gives

$$R_I = \exp\left[\frac{-m}{\tilde{n}[\psi(\tilde{n}+T)-\psi(\tilde{n})]}\right] \tag{40}$$

and

$$\sum_{i=1}^{m}\frac{\sum_{j=0}^{n_{i,2}}\left(\prod_{q=1}^{n_{i,1}}\frac{1}{\tilde{n}+v_i+\tau_{i,j}-q}\right)\sum_{q=1}^{n_{i,1}}\frac{1}{\tilde{n}+v_i+\tau_{i,j}-q}}{\sum_{j=0}^{n_{i,2}}\left(\prod_{q=1}^{n_{i,1}}\frac{1}{\tilde{n}+v_i+\tau_{i,j}-q}\right)} = \mathrm{m}\frac{[\psi'(\tilde{n})-\psi'(\tilde{n}+T)]}{\psi(\tilde{n}+T)-\psi(\tilde{n})} \tag{41}$$

The empirical Bayes estimate of initial reliability can now be found by using Equation (41) to solve the unknown value $\tilde{n}$ and then substituting this value into Equations (40).

3.8 *Simulation Analysis of Model Adequacy*

To examine the usefulness of the proposed approach for reliability projection, failure and corrective action simulations were conducted on hypothetical complex discrete use systems under various test circumstances. Simulation analysis enabled the calculation of the system's true initial reliability upon entering the test. In addition, the FEF and MS management planning metric distributions were identified using set mean and variance values to permit greater control of the simulation within known bounds. Consequently, the prior and posterior reliability estimates could then be calculated at the end of each simulation series and compared to the hypothetical system known true prior and posterior values.

Simulation techniques were utilized to permit the assessment of various models including:

- Hall's Bayesian approach for a known k number of possible system failure modes,
- Hall's Bayesian approach for a system containing an infinite number of failure modes, and
- the proposed Bayesian approach.

All three reliability projection approaches were assessed by simulation in the Python general-purpose programming language, with each approach measured against the same simulated data to ensure accurate comparisons of performance and results.

3.8.1 Mathematical Model, Code and Simulation Verification

Before conducting any detailed modeling and simulation, it is accepted good practice to undertake verification activities that confirm the model and simulation implementations. Verification includes confirming that the associated data outputs represent the conceptual model description and are accurate representations from the perspective of their intended use [44].

Prior to conducting any detailed empirical Bayes estimate simulation, the mathematical model for estimation of the empirical Bayes initial reliability estimates (Equation 40), together with the supporting mathematical equation used to solve the unknown value $\tilde{n}$ were verified via hand calculations to confirm that the output values were satisfactory and within reasonably expected bounds. This was performed to ensure that mathematical errors had been eliminated that may have occurred and not been detected during initial model development and that the model was providing the results expected.

Verification included calculating the values derived for $\tilde{n}$ and R_I using ten randomly generated failure probabilities from a known Beta(a,b) distribution and assuming each failure probability resulted in five single occurrences of observed failures and five double occurrences of observed failures for each failure probability.

It was then further assumed that two of each of the single and double observed failure modes (four failure modes in total) were subjected to corrective actions during the test phase, which was set at 50 trials. The remaining failure modes were then assumed to be addressed at the end of the current test phase (at $v_i = T$).

One of each of the double and single failure modes corrected was then assumed to result in an additionally observed failure ten trial demands after the corrective action implementation (i.e., at $v_i + 10$).

The fix effectiveness factor for all corrective actions was set at 0.8 for simplicity.

Hand calculations were then completed using the above process several times using slightly different input variables each time to arrive at solutions for the $\tilde{n}$ and R_I values.

These values were then entered into the developed Python code as coupled matrices to assess the output compared to the hand calculations.

Finally, the process was repeated to assess the proposed model reliability expectation (Equation 21) and variance (Equation 26). Again, hand calculations were completed and then compared to the Python code results. Only when these values were identical and any differences investigated and rectified could the developed models and simulation code be assured as suitable for simulation and evaluation of outcomes.

3.8.2 Simulation Outcomes

Table 1 outlines the results of five selected simulated failures and corrective actions during a single reliability growth test phase. One thousand iterations for each tended to provide a good balance between projection solution fidelity, accuracy, and computational resource usage in the cases examined. A greater number of computational iterations produced no significant change in the overall result obtained but consumed a more considerable computational resource.

Table 1: Simulation input variables.

Case	No. Failure Modes k	No. Test Demands T	Mode Failure Prob Mean $p(p_i)_\mu$	Mode Failure Prob Var $p(p_i)_{\sigma^2}$	FEF Mean d_μ	FEF Var d_{σ^2}	Prob Fix During Test $p(fix)$
1	50	200	0.005	0.00007	0.7	0.01	0.3
2	50	200	0.002	0.00002	0.7	0.01	0.3
3	50	400	0.003	0.00005	0.7	0.01	0.3
4	100	300	0.002	0.00005	0.7	0.01	0.5
5	100	200	0.003	0.00005	0.7	0.01	0.4

Table 2 displays the results of the selected reliability projection simulation cases outlined at a two-sided 80% confidence level. For initial estimator/true and posterior/true case each case, three reliability metrics were compared, including the lower confidence limit, the mean, and the upper confidence limit (separated in Table 2 by individual obliques). The absolute relative error between the estimator and true means was also calculated.

In all instances, the Bayes prior reliability estimator and the posterior estimate results provide a reasonable indicator of the system's true reliability.

Table 2: Simulation true reliability versus Bayesian estimates comparison.

Case	No. Obs Fails / Modes n/m	True Initial Reliability $R_{I,true}$	Initial Reliability Estimator $R_{I,est}$	Initial Estimator Mean Relative Error R_I $Rel\ Err$	True Posterior Reliability $R_{P,true}$	Posterior Reliability Estimator $R_{P,est}$	Posterior Estimator Mean Relative Error R_P $Rel\ Err$
1	20/18	0.72/0.78/0.84	0.71/0.79/0.87	0.02	0.88/0.90/0.93	0.94/0.95/0.97	0.05
2	12/10	0.87/0.90/0.94	0.83/0.88/0.94	0.02	0.94/0.95/0.97	0.96/0.97/0.98	0.02
3	23/15	0.80/0.86/0.92	0.80/0.86/0.92	0.00	0.93/0.95/0.97	0.97/0.98/0.99	0.03
4	19/16	0.74/0.82/0.90	0.76/0.83/0.90	0.01	0.90/0.93/0.96	0.97/0.98/0.99	0.05
5	25/23	0.67/0.74/0.81	0.67/0.75/0.83	0.02	0.85/0.88/0.91	0.93/0.94/0.96	0.06

Figure 3 demonstrates a comparison between the true and estimator initial distributions graphically for Case 1.

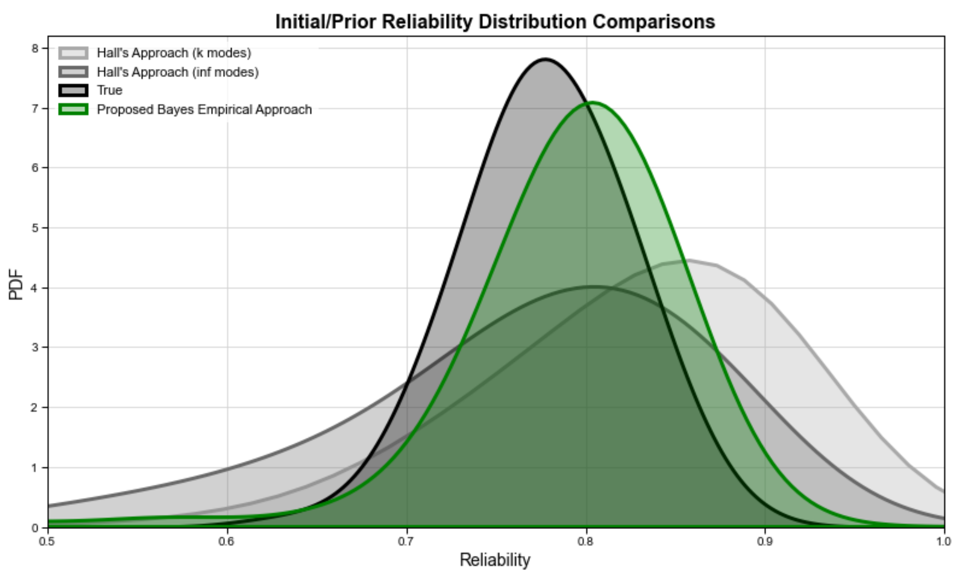


Figure 3. Case 1 initial reliability true and estimator distribution comparisons.

Figure 4 displays a comparison between the true and estimator posterior distributions graphically for Case 1.

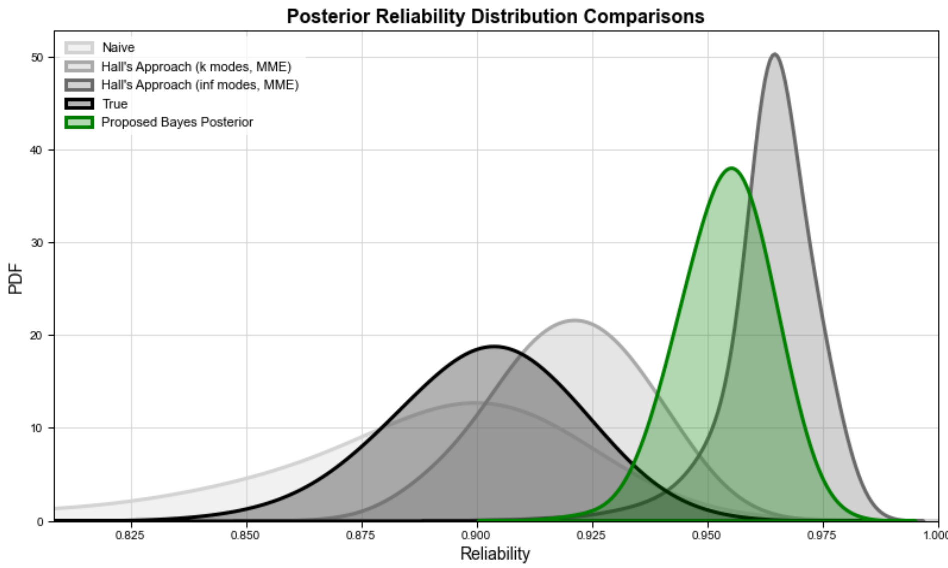


Figure 4. Case 1 posterior reliability true and estimator distribution comparisons.

The mean relative error values for both the Bayesian technique proposed here and Hall's approach for both k and infinite failure modes provide a useful measure of performance of all models when compared against the true system initial and post-test phase reliability. The absolute relative error for the simulation cases is defined as:

$$\text{Relative Error} = \frac{\left|\hat{R} - R\right|}{R} \tag{42}$$

where R is the true system reliability, either initial or post-test phase after all corrective actions have been implemented, and $\hat{R}$ is the model estimate resulting from the simulated data. It is important to note that the use of absolute relative error in this manner may provide a conservative indication of the performance of the estimators. As the reliability estimators are small numerical values, seemingly minor differences may actually be large percentage values and result in a high relative error.

Figure 5 displays the Case 1 simulation initial reliability estimators cumulative relative errors for model performance comparison.

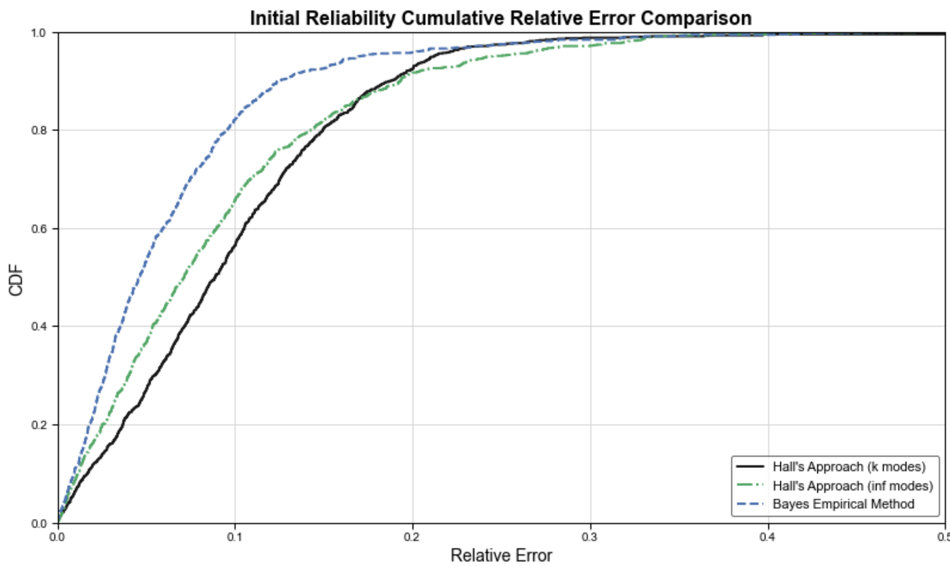


Figure 5. Case 1 initial reliability estimator cumulative relative error comparisons.

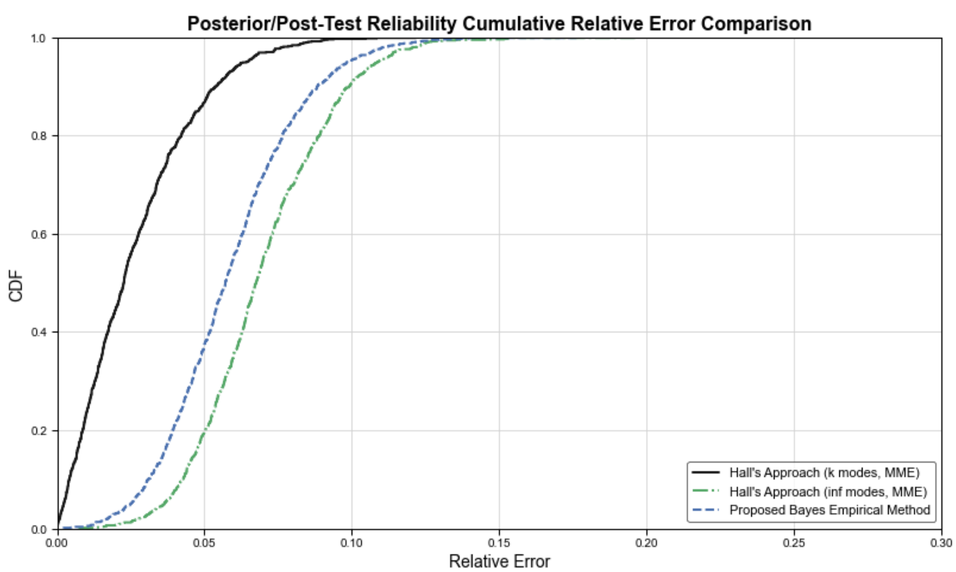


Figure 6. Case 1 posterior reliability estimator cumulative relative error comparisons.

Annex A to Chapter 3 includes the distribution and cumulative relative error plots for the remaining four simulation cases.

3.9 *Discussion*

Table 1 and Table 2 demonstrate the utility of the proposed method through the selected demonstration cases. Generally, both the proposed initial and posterior reliability estimators represent the true system-level reliability well. However, the power of the proposed approach is displayed in Figure 3 to Figure 6.

Figure 3 shows the prior distributions produced through simulation for three approaches, Hall's k modes [35] [36], Hall's infinite number of modes [35] [36] and the proposed empirical Bayes approach together with the true initial reliability distribution. From this plot, it can be observed that the proposed approach best represents the true initial reliability compared to the other methods. Moreover, this is true for a large proportion of the simulations conducted using different input variables.

Figure 4 displays the posterior distributions for both of Hall's approaches (k and infinite), the naive reliability distribution (reliability estimation considering only observed failures), the proposed Bayesian approach, and the true post-test reliability distribution after all corrective actions were implemented. This figure shows that the naive reliability distribution spread, based on a simple point estimate of reliability from observed failures, is excessive and unsuitable for reliability estimation. The proposed Bayesian approach performs better than Hall's approach utilizing an infinite number of modes, but not Hall's model for a known number of k modes. This is to be expected as, in reality, the number of failure modes in a complex system remains unknown, so any model that relies on knowing this detail precisely is of limited utility. More direct comparisons are possible but unrealistic, given that the number of modes must be known.

Figure 5 displays the initial true and prior estimator distributions cumulative relative error comparisons for Case 1. This image demonstrates the superior performance of the proposed empirical Bayes approach over both Hall's k and infinite number of modes prior reliability estimators.

Figure 6 demonstrates the cumulative relative error for the posterior distributions similarly to Figure 5, demonstrating the same for the initial reliability estimator distributions. In this case, Hall's model for k known failure modes performs significantly better than the others; however, as highlighted, it is not typical that the total number of failure modes would be known in reality and a model of this nature has only limited utility. Figure 6 also displays the performance of the proposed Bayesian approach as an improvement over Hall's infinite modes model.

These cumulative relative error plots highlight the benefits of the proposed approach over Hall's method.

3.10 *Reliability Growth Planning Expansion*

The proposed Bayesian projection model has been demonstrated to provide improvements in estimation accuracy when compared to the approaches developed by

Hall et al. [30] [31] [33]. An approach has been developed that permits the same mathematical modeling efforts to be used for reliability growth planning purposes that produce more realistic plans than previous methods and are flexible enough to enable quick adaption to changing circumstances or unexpected reliability growth program outcomes.

The planning approach extension follows the projection methodology described in this chapter but combines individual test events or phases into a cohesive reliability program plan. Figure 7 demonstrates the concept diagrammatically.

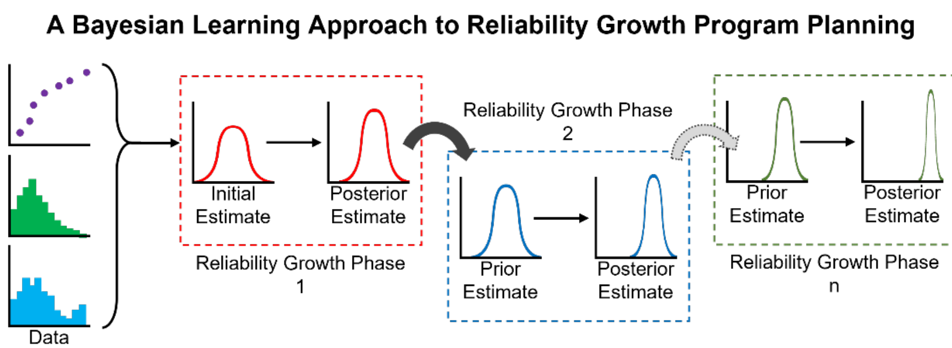


Figure 7. Reliability growth program planning Bayesian learning concept.

In the preliminary stages of planning for a system-level reliability growth program, data from a wide range of sources may be used to provide an initial estimate of reliability on test entry. Potential data sources could include previous testing, manufacturer's component, subsystem or system testing, expert opinion, similar systems or data gained from prototypes and earlier designs.

The proposed Bayesian reliability projection model is then developed virtually, and simulation approaches are used to produce system failure and corrective action data.

Over many iterations, the failure and corrective action data are used to develop the posterior estimate of reliability on test completion and after all corrective actions have been implemented. The number of corrective actions implemented depends on the simulated observed failures and the expected management strategy. Various outcomes can be assessed for a single test phase by repeating the process with variables that reflect different resource allocations, management strategies, or the effectiveness of corrective actions.

For each reliability growth test phase, the posterior reliability distribution for the immediate past phase becomes the prior reliability distribution for the next. This process can then be repeated, as demonstrated in Figure 7, for as many test phases as needed.

Chapter 4 compares empirical Bayes methods for estimating initial system reliability from test data and expands on the methods described in this chapter.

Chapter 5 demonstrates the utility of the proposed Bayesian projection approach through the development of a system-level reliability growth plan. The plan established extends the proposed combined projection and simulation method to produce projection data from developmental, demonstration, and operational testing. The reliability program plan created can easily be extended to project reliability at any part of the future system life cycle.

3.11 *Additional Reliability Metrics*

Several useful reliability projection and growth planning metrics have also been developed through the simulation of failures and corrective actions using the proposed Bayesian approach. These include:

- expected number of failure modes observed during the test,
- estimated cumulative number of observed failure modes,
- Number of total failures likely to be observed before and after the implementation of corrective actions, and
- number of observed modes likely and not likely to be subjected to corrective action during the test.

Each of these metrics may be described via their expectation value and confidence bounds due to the simulations carried out. Example plots of each of these metrics can be found in sample Python code within Appendix 2.

Further reliability metrics may also be developed from the simulation code that describe the influences of various parameters on the test phase final reliability estimate.

3.12 *Reliability Growth Potential Estimation*

Reliability growth potential is the theoretical upper limit on the reliability of a system that can be achieved by finding and mitigating the proportion of observed failure modes in accordance with the management strategy at a specified level of fix effectiveness. It is defined as

$$\bar{R}_{GP} = \mathrm{R}_I^{1-\overline{MS}\cdot\mu_d} \tag{43}$$

where $\bar{R}_{GP}$ is the reliability growth expectation, R_I is the initial reliability, $\overline{MS}$ is the mean expected management strategy, and μ_d is the expected mean fix effectiveness.

While it is possible to add confidence limits to the reliability growth potential, this is only useful in demonstrating the influence of the three critical unknown reliability growth planning factors (initial reliability, mean management strategy, and mean fix effectiveness) to decision makers that might be prone to adjusting their management goals during test activities. A simple graphic demonstrating the upper and lower confidence limits on the reliability growth potential across a range of management strategies is helpful in demonstrating the influence management decisions have on test outcomes.

3.13 *Comparison Plots Simulation versus Mathematical Model*

In addition to the additional reliability metrics described within Section 3.11, three additional plots have also been identified that are useful in comparing the model to the simulated outcome results. These include:

- comparison of simulation and model mode failure prior probability distributions,

- comparison of simulation posterior and model mode failure prior probability distributions, and

- comparison of simulation and model fix effectiveness factor distributions.

Further comparison plots may also be developed from the simulation code that compare the mathematical model inputs with the simulation iteration outputs.

3.14 *Model Goodness-of-Fit*

The model goodness-of-fit may be graphically assessed by plotting the cumulative number of observed failure modes against the estimate of the cumulative number of expected observed failure modes given by

$$\hat{\mu}_t = \left(\frac{m}{\sum_{j=0}^{T-1} \frac{1}{\tilde{n} + j}} \right) \cdot \left(\frac{\Gamma'(\tilde{n} + t)}{\Gamma(\tilde{n} + t)} - \frac{\Gamma'(\tilde{n})}{\Gamma(\tilde{n})} \right) \tag{44}$$

where $\hat{\mu}_t$ is the estimate of the expected number of observed failure modes at test demand t.

3.15 *Conclusions*

The presented method for constructing a real-world representative reliability growth projection model for complex discrete-use systems has been demonstrated as superior to current approaches under many different circumstances. The mixing of Bayesian techniques and simulation methods provides a reasonable estimation of the likely outcomes of reliability growth test phases when the proposed approach is used for planning. The approach offered considers reliability growth planning uncertainties resulting in a practical and real-world representative model. The approach models the uncertainties in the underlying failure mode probability, corrective action effectiveness, and the management strategy for dealing with observed failure modes within a Bayesian framework. Uncertainty quantification is essential in ensuring that decision makers know and understand the range of potential consequences and the breadth of risk of various management strategies and corrective action schemes when conducting reliability growth testing activities.

Annex A to Chapter 3: Selected Simulation Results

A.3.1 Introduction

The purpose of this annex is to present the results for several selected simulations. The simulations presented are the results of Case 2 to Case 5 within Table 1.

A.3.2 Simulation Case 2

A.3.2.1 Input Variables

Table 3: Simulation Case 2 input variables.

Case	No. Failure Modes k	No. Test Demands T	Mode Failure Prob Mean $p(p_i)_\mu$	Mode Failure Prob Var $p(p_i)_{\sigma^2}$	FEF Mean d_μ	FEF Var d_{σ^2}	Prob Fix During Test $p(fix)$
2	50	200	0.002	0.00002	0.7	0.01	0.3

A.3.2.2 Tabulated Results

Table 4: Simulation Case 2 result comparison.

Case	No. Obs Fails / Modes n/m	True Initial Reliability $R_{I,true}$	Initial Reliability Estimator $R_{I,est}$	Initial Estimator Mean Relative Error $R_I\ Rel\ Err$	True Posterior Reliability $R_{P,true}$	Posterior Reliability Estimator $R_{P,est}$	Posterior Estimator Mean Relative Error $R_P\ Rel\ Err$
2	12/10	0.87/0.90/0.94	0.83/0.88/0.94	0.02	0.94/0.95/0.97	0.96/0.97/0.98	0.02

A.3.2.3 Initial/Prior Distribution Plot Comparison

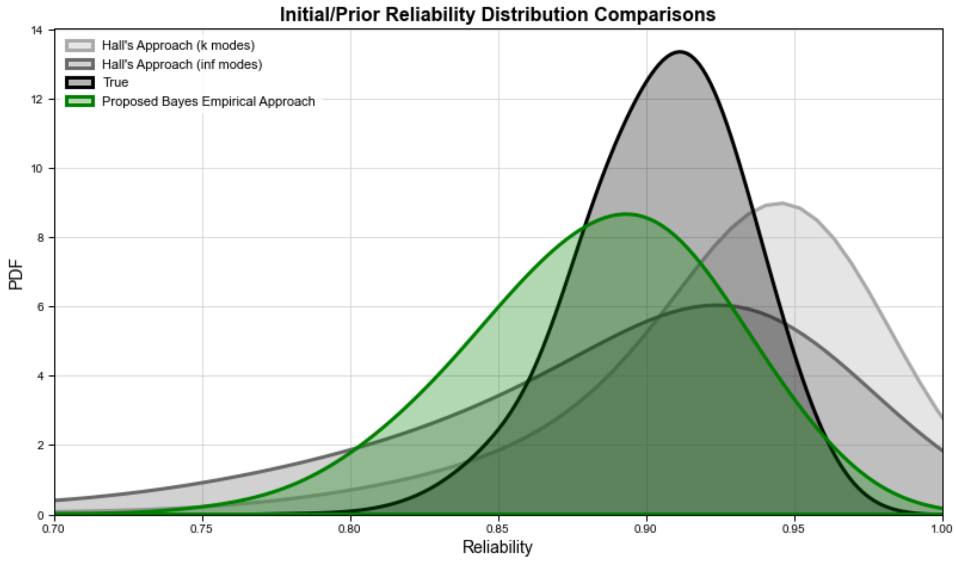


Figure 8. Case 2 initial/prior distribution comparison plot.

A.3.2.4 Post-Test/Posterior Distribution Plot Comparison

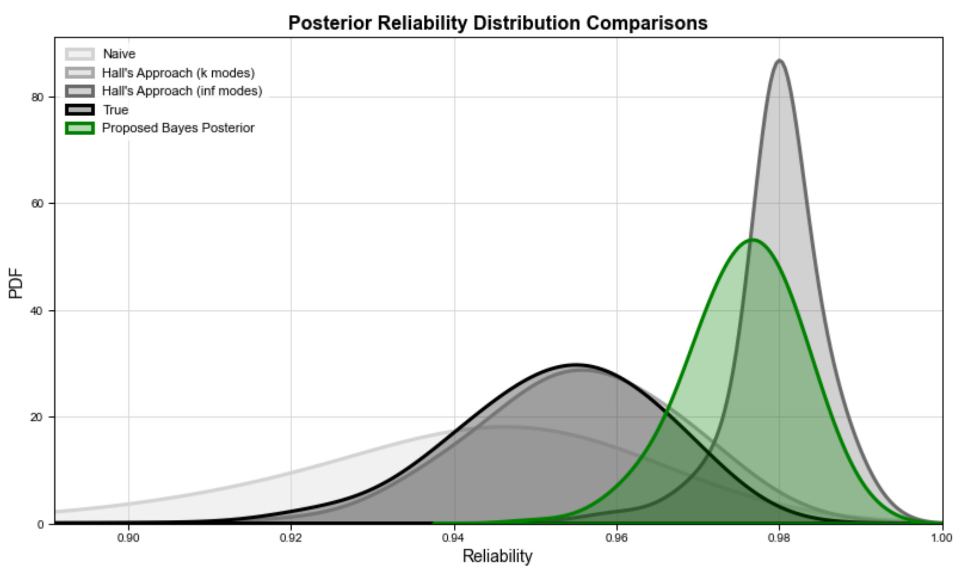


Figure 9. Case 2 post-test/posterior distribution comparison plot.

A.3.2.5 Initial Reliability Estimator Relative Error Comparison

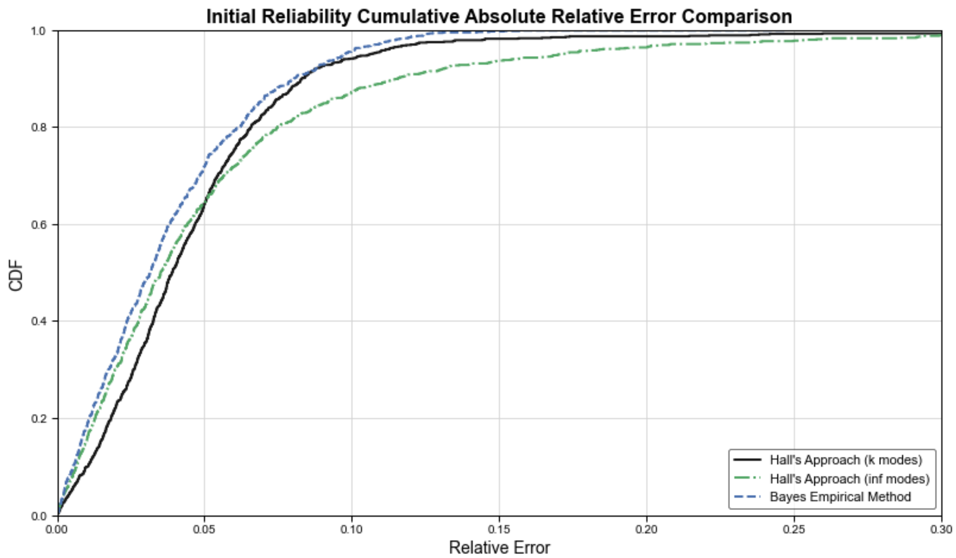


Figure 10. Case 2 initial reliability estimator relative error comparison.

A.3.2.6 Posterior Reliability Estimator Relative Error Comparison

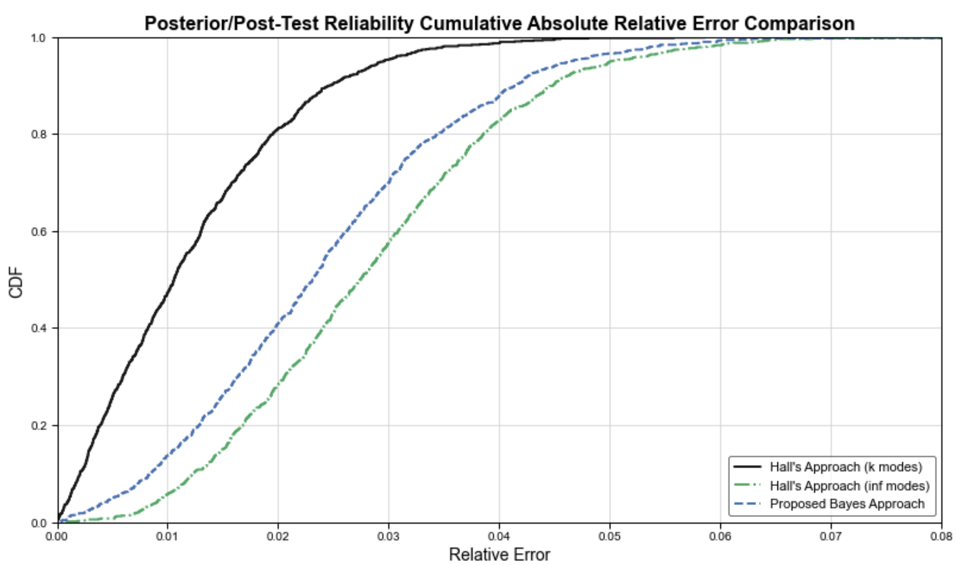


Figure 11. Case 2 posterior reliability estimator relative error comparison.

A.3.2.7 Observation Summary

The results of the Case 2 simulation over 1,000 iterations demonstrate that the developed empirical Bayes initial reliability estimator performs relatively well when compared against Hall's approach for an infinite number of failure modes. Based on an 80% two-sided confidence limit, the cumulative mean relative error in the estimator is only about 2%.

In terms of distribution coverage, Figure 8 demonstrates that the proposed empirical Bayes estimator has superior coverage when compared against Hall's k and infinite mode approaches. In this case, the proposed empirical Bayes estimator mean is on the conservative side of the true system initial reliability. In contrast, the Hall infinite approach estimator tends to be more optimistic.

Figure 10 conveys the cumulative relative error comparison between the proposed Bayes empirical approach and Hall's methods. This plot demonstrates the utility of the proposed estimator under the Case 2 range of variables and indicates superior performance when the estimators are compared against the true initial reliability.

The performance of the posterior distributions is evident in Figure 9. Hall's k modes approach performs best within the variable constraints of Case 2; however, as previously outlined in Chapter 3, it is improbable that the true number of failure modes is ever known for a complex system. Consequently, the ability to utilize Hall's k approach in assessing a real-world system is limited. Hall's k mode approach is included for completeness and idealistic comparisons in all simulation case studies presented within this thesis.

The estimator's posterior cumulative mean relative error is about 2%, similar to the initial reliability estimator error. Figure 11 displays that the proposed posterior reliability estimator performs marginally better than Hall's for an infinite number of failure modes.

Overall, the proposed empirical Bayes initial reliability approach and the proposed Bayesian posterior estimates perform better than Hall's approach under the variable constraints of simulation Case 2.

A.3.3 Simulation Case 3

A.3.3.1 Input Variables

Table 5: Simulation Case 3 Input Variables.

Case	No. Failure Modes k	No. Test Demands T	Mode Failure Prob Mean $p(p_i)_\mu$	Mode Failure Prob Var $p(p_i)_{\sigma^2}$	FEF Mean d_μ	FEF Var d_{σ^2}	Prob Fix During Test $p(fix)$
3	50	400	0.003	0.00005	0.7	0.01	0.3

A.3.3.2 Tabulated Results

Table 6: Simulation Case 3 Result Comparison.

Case	No. Obs Fails / Modes n/m	True Initial Reliability $R_{I,true}$	Initial Reliability Estimator $R_{I,est}$	Initial Estimator Mean Relative Error $R_I\ Rel\ Err$	True Posterior Reliability $R_{P,true}$	Posterior Reliability Estimator $R_{P,est}$	Posterior Estimator Mean Relative Error $R_P\ Rel\ Err$
3	23/15	0.80/0.86/0.92	0.80/0.86/0.92	0.00	0.93/0.95/0.97	0.97/0.98/0.99	0.03

A.3.3.3 Initial/Prior Distribution Plot Comparison

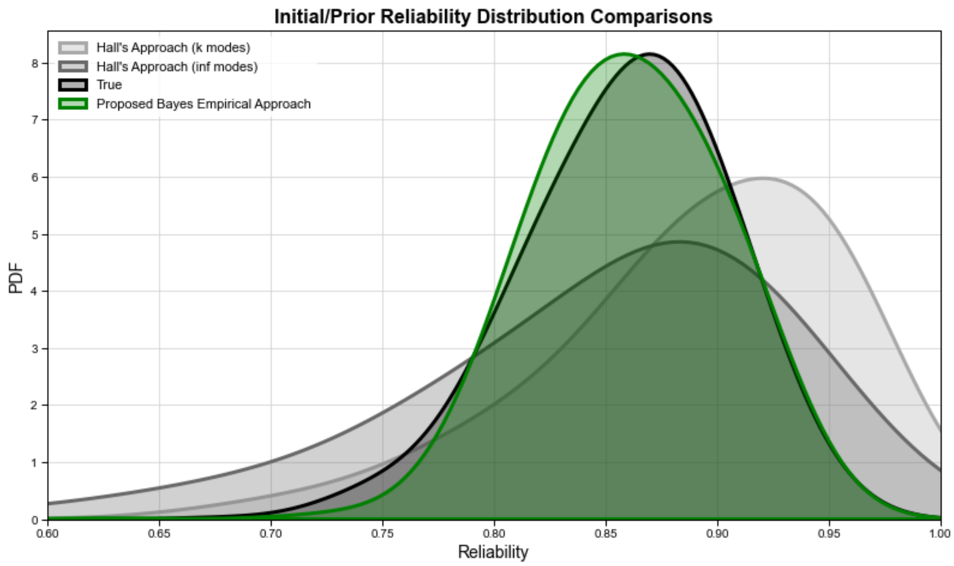


Figure 12. Case 3 initial/prior distribution comparison plot.

A.3.3.4 Post-Test/Posterior Distribution Plot Comparison

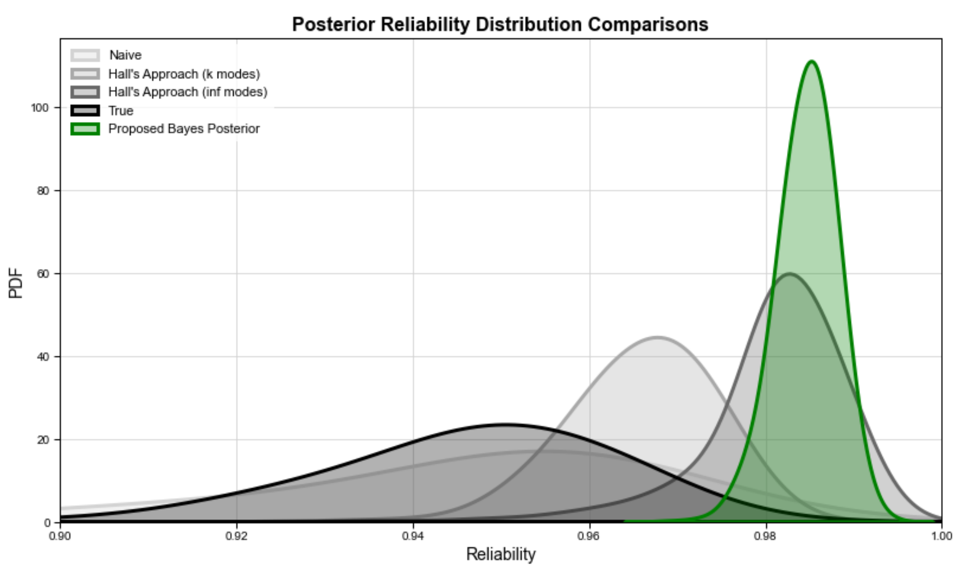


Figure 13. Case 3 post-test/posterior distribution comparison plot.

A.3.3.5 Initial Reliability Estimator Relative Error Comparison

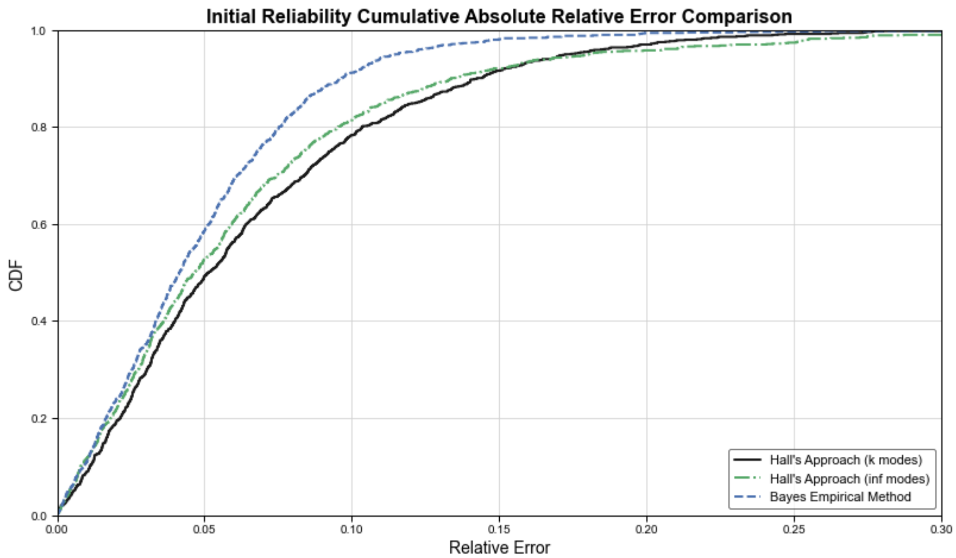


Figure 14. Case 3 initial reliability estimator relative error comparison.

A.3.3.6 Posterior Reliability Estimator Relative Error Comparison

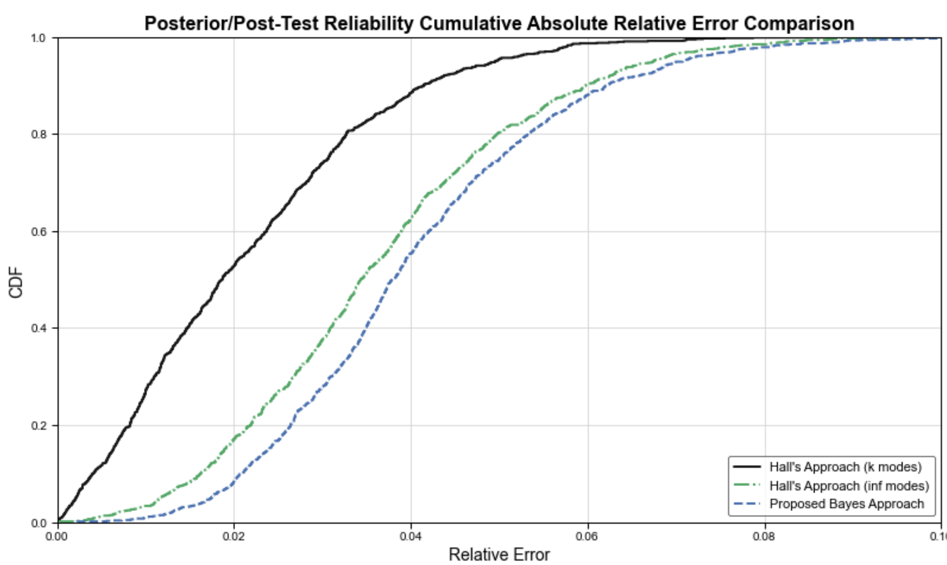


Figure 15. Case 3 posterior reliability estimator relative error comparison.

A.3.3.7 Observation Summary

The results of the Case 3 simulation demonstrate that the developed empirical Bayes initial reliability estimator performs very well when compared against Hall's approach for an infinite number of failure modes. The cumulative mean relative error in the estimator is zero, and the empirical Bayes estimator distribution achieves 100% coverage over the true initial reliability distribution. The exception distribution coverage is clearly displayed in Figure 12.

Figure 10 demonstrates the cumulative relative error comparison between the proposed Bayes empirical approach and Hall's methods, with the proposed approach offering better initial reliability estimation outcomes.

The performance of the posterior distributions is evident in Figure 13. Under the variable constraints for the Case 3 simulation, Hall's infinite approach proves marginally superior. Note that the proposed Bayesian and Hall's infinite approaches tend to over-optimistically estimate the true system posterior reliability by approximately 4%. Both posterior estimate distributions also have inadequate coverage of the true system posterior reliability distribution.

Figure 15 displays the significant mean relative error in all approaches considered in simulation Case 3.

Overall, the proposed empirical Bayes initial reliability estimator approach performs very well under simulation Case 3. Neither Hall's nor the proposed posterior reliability estimator approaches estimate the true posterior reliability particularly well, yet both remain suitable for reliability growth planning purposes.

A.3.4 Simulation Case 4

A.3.4.1 Input Variables

Table 7: Simulation Case 4 input variables.

Case	No. Failure Modes k	No. Test Demands T	Mode Failure Prob Mean $p(p_i)_{\mu}$	Mode Failure Prob Var $p(p_i)_{\sigma^2}$	FEF Mean d_{μ}	FEF Var d_{σ^2}	Prob Fix During Test $p(fix)$
4	100	300	0.002	0.00005	0.7	0.01	0.5

A.3.4.2 Tabulated Results

Table 8: Simulation Case 4 result comparison.

Case	No. Obs Fails / Modes n/m	True Initial Reliability $R_{I,true}$	Initial Reliability Estimator $R_{I,est}$	Initial Estimator Mean Relative Error R_I $Rel\ Err$	True Posterior Reliability $R_{P,true}$	Posterior Reliability Estimator $R_{P,est}$	Posterior Estimator Mean Relative Error R_P $Rel\ Err$
4	19/16	0.74/0.82/0.90	0.76/0.83/0.90	0.01	0.90/0.93/0.96	0.97/0.98/0.99	0.05

A.3.4.3 Initial/Prior Distribution Plot Comparison

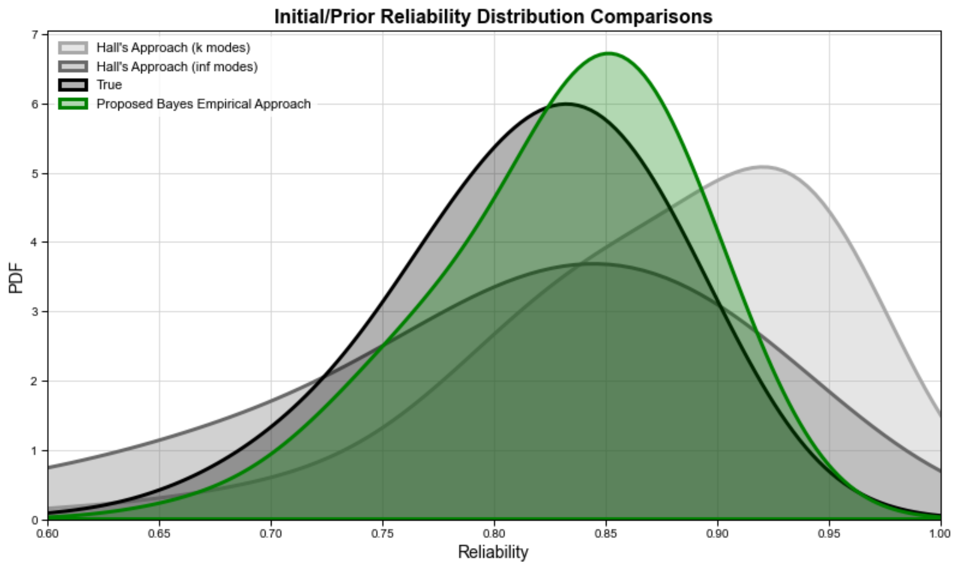


Figure 16. Case 4 initial/prior distribution comparison plot.

A.3.4.4. Post-Test/Posterior Distribution Plot Comparison

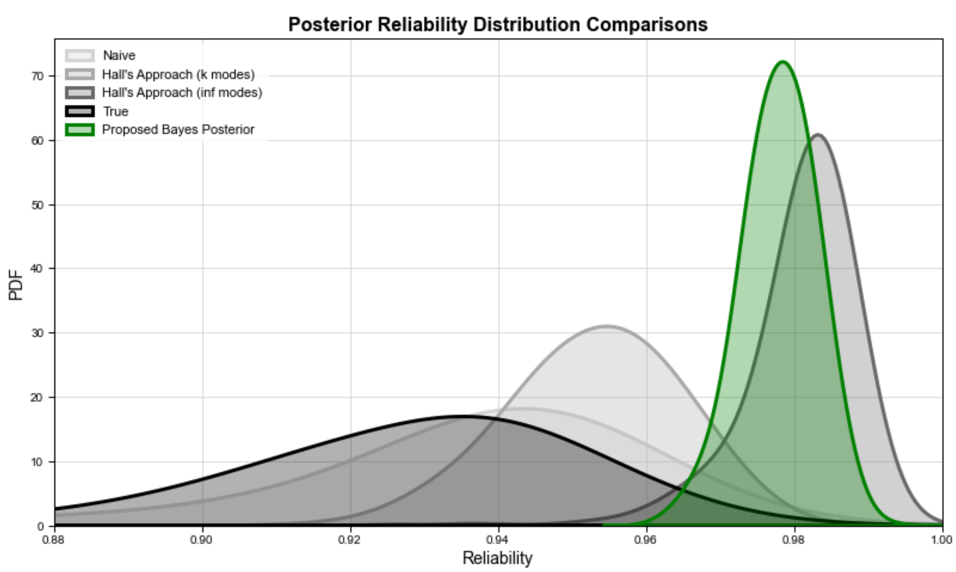


Figure 17. Case 4 post-test/posterior distribution comparison plot.

A.3.4.5 Initial Reliability Estimator Relative Error Comparison

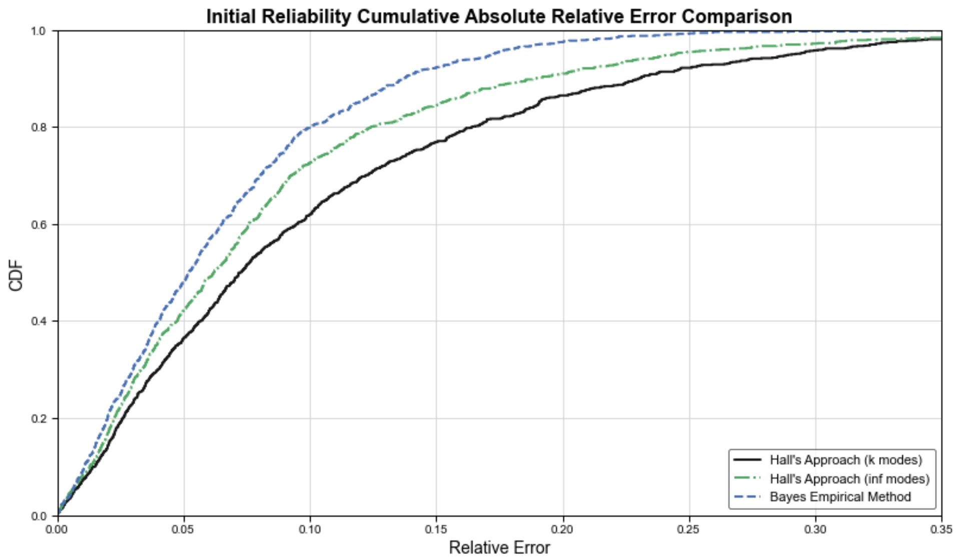


Figure 18. Case 4 initial reliability estimator relative error comparison.

A.3.4.6 Posterior Reliability Estimator Relative Error Comparison

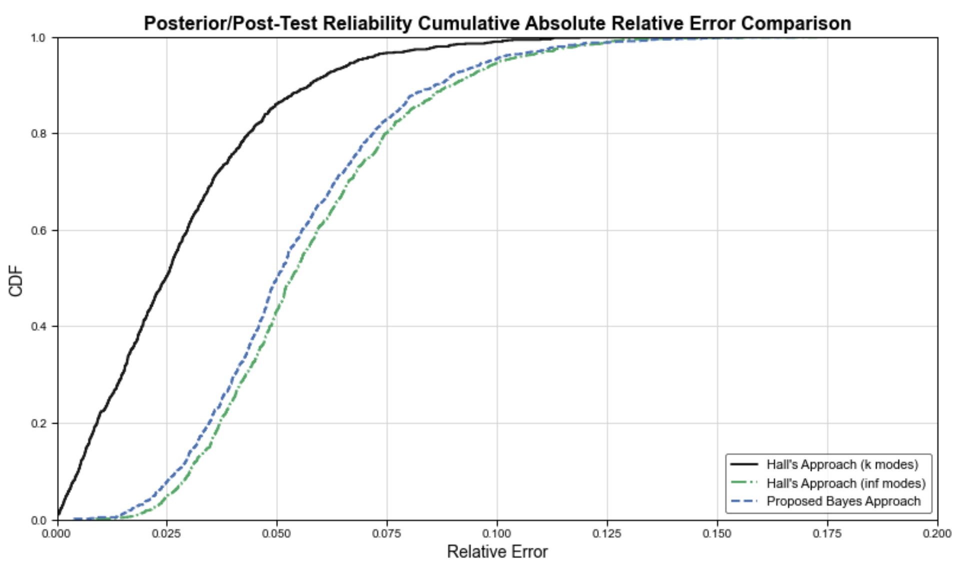


Figure 19. Case 4 posterior reliability estimator relative error comparison.

A.3.4.7 Observation Summary

Simulation Case 4 demonstrates similar estimator outcomes when compared to simulation Case 3.

In terms of estimation of initial system-level reliability, all approaches provide satisfactory results and reasonable true distribution coverage.

The posterior distribution comparison within Figure 13 is somewhat similar to the Case 3 posterior distribution comparison demonstrated in Figure 9. However, in this case, the proposed approach offers marginal improvements compared to Hall's infinite modes methodology.

Importantly, note again that both the proposed Bayesian and Hall's infinite approaches tend to over-optimistically estimate the true system posterior reliability by approximately 3-4%. Both posterior estimate distributions also again have inadequate coverage of the true system posterior reliability distribution.

Overall, the proposed empirical Bayes initial reliability estimator approach performs very well under simulation Case 4. Neither Hall's nor the proposed posterior reliability estimator approaches estimate the true posterior reliability well. However, both are accurate enough to remain suitable for reliability growth planning purposes.

A.3.5 Simulation Case 5

A.3.5.1 Input Variables

Table 9: Simulation Case 5 input variables.

Case	No. Failure Modes k	No. Test Demands T	Mode Failure Prob Mean $p(p_i)_\mu$	Mode Failure Prob Var $p(p_i)_{\sigma^2}$	FEF Mean d_μ	FEF Var d_{σ^2}	Prob Fix During Test $p(fix)$
5	100	200	0.003	0.00005	0.7	0.01	0.4

A.3.5.2 Tabulated Results

Table 10: Simulation Case 5 result comparison.

Case	No. Obs Fails / Modes n/m	True Initial Reliability $R_{I,true}$	Initial Reliability Estimator $R_{I,est}$	Initial Estimator Mean Relative Error $R_I\ Rel\ Err$	True Posterior Reliability $R_{P,true}$	Posterior Reliability Estimator $R_{P,est}$	Posterior Estimator Mean Relative Error $R_P\ Rel\ Err$
5	25/23	0.67/0.74/0.81	0.67/0.75/0.83	0.02	0.85/0.88/0.91	0.93/0.94/0.96	0.06

A.3.5.3 Initial/Prior Distribution Plot Comparison

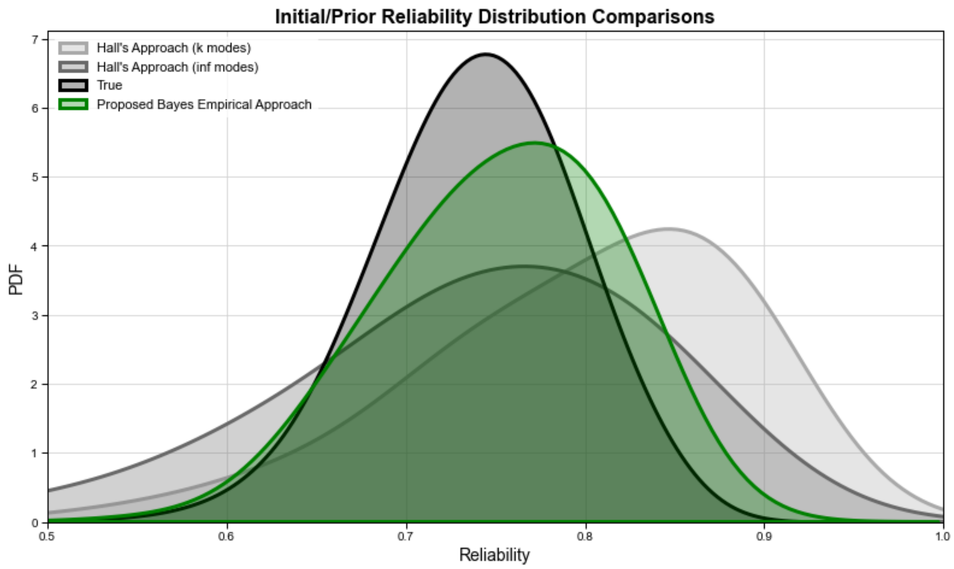


Figure 20. Case 5 initial/prior distribution comparison plot.

A.3.5.4 Post-Test/Posterior Distribution Plot Comparison

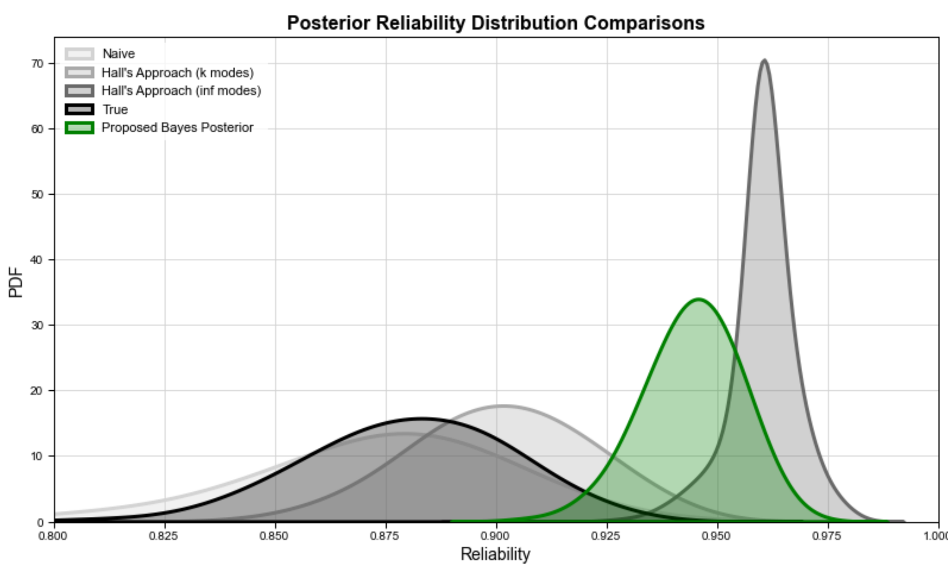


Figure 21. Case 5 post-test/posterior distribution comparison plot.

A.3.5.5 Initial Reliability Estimator Relative Error Comparison

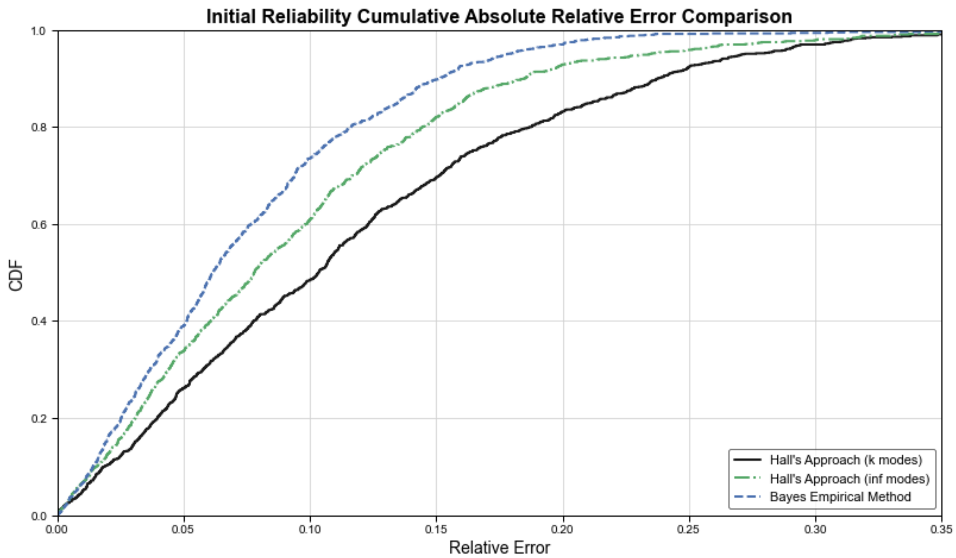


Figure 22. Case 5 initial reliability estimator relative error comparison.

A.3.5.6 Posterior Reliability Estimator Relative Error Comparison

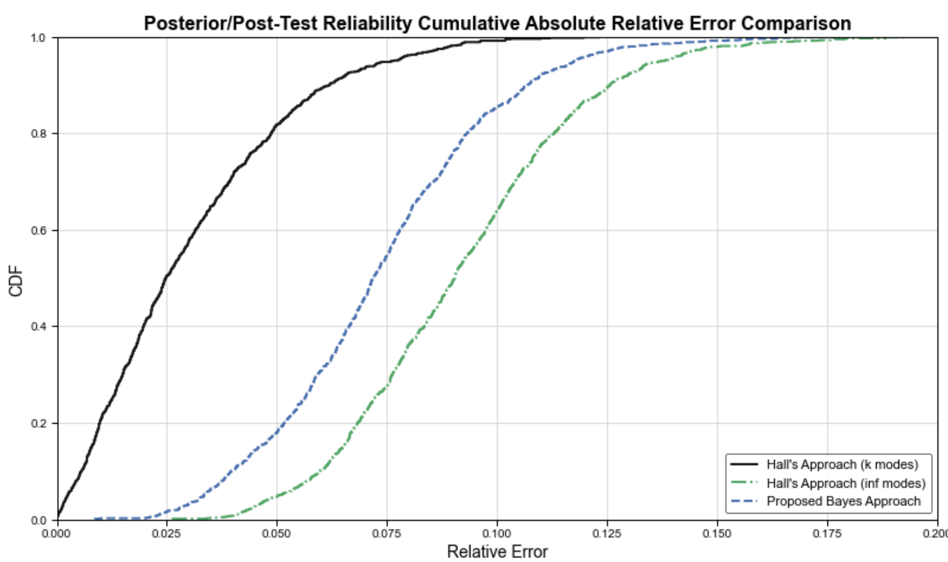


Figure 23. Case 5 posterior reliability estimator relative error comparison.

A.3.5.7 Observation Summary

Simulation Case 5 demonstrates that all approaches provide satisfactory results in terms of estimation of initial system-level reliability. Note that of the approaches considered that the proposed empirical Bayes estimator offered has superior distribution coverage over the true distribution.

Figure 21 and Figure 23 demonstrate that the proposed Bayesian posterior estimator approach offers improvements over Hall's infinite modes approach. Intuitively, we expect this to be the case as Hall's approach only considers the first observation of each failure mode (or "first occurrence on test" (FOT) to use Hall's terminology). In contrast, the proposed approach utilizes a broader range of available data, including all observed mode failure occurrences on test and when corrective actions are likely to be implemented.

Chapter 4: A Comparison of Empirical Bayes Hyperparameter Approaches for Discrete-Use System Initial Reliability Estimation[2]

4.1 *Introduction*

Bayesian probability is a statistic theory based on the Bayesian interpretation of probability, where probability expresses a degree of belief in an event [45]. The degree of trust that a person holds may rise from prior knowledge about the event [46], such as the results of previous experiments or tests, or a quantified personal belief [47]. Bayesian probability differs from another interpretation of probability, the frequentist interpretation, which considers probability as the limit of an event's relative frequency after many trials [48].

The empirical Bayes method is the collective terminology used to denote statistical inference procedures in which the data developed during a reliability growth program is used to estimate the initial reliability prior distribution. This approach differs from standard Bayesian methods, for which the prior distribution is typically fixed before any failure modes are observed. The use of a fixed prior may lead to significant concerns if an erroneous distribution is selected through the use of incomparable

---

[2] The content of this chapter was presented at the Australian Integrated Project Engineering Congress (IPEC) May 26-28, 2021. Note that the presented paper was tailored to an audience that included those with a limited reliability engineering background. Consequently, some concepts are explained in more detail than would be necessary for a knowledgeable audience.

historical data or invalid expert opinion. Despite this difference in perspective, empirical Bayes may be viewed as an approximation to a fully Bayesian treatment of a hierarchical model wherein the parameters at the highest level of the hierarchy are set to their most likely values instead of being integrated out.

This chapter will describe and assess two approaches for determining the initial reliability of a discrete-use system tested under constraints utilizing only the data derived from the actual testing. The importance of initial reliability estimation based on evidence cannot be understated.

4.2 *Empirical Bayes Methods*

Bayesian approaches offer unique insights into reliability test outcomes that may not be obvious when other traditional techniques are adopted. The Bayesian approach uses previous experience and new test data combined when applying statistical tools to assess reliability metrics [49]. A posterior distribution is derived from a prior distribution and a likelihood function in the Bayesian approach, and any following tests are conducted using the derived posterior distribution. When new sample data is added, this posterior distribution is then employed as a prior distribution in the process of producing a further posterior distribution. This cyclical use of the posterior distribution as a prior distribution in reliability testing of a finite population is known as the Bayesian learning process [50]. However, looking backward can be problematic at the commencement of a reliability growth test sequence to establish an initial estimate of system reliability that moves beyond previous experience or subjective judgment. In selecting an estimate for initial system reliability based on limited early test data, empirical Bayes approaches may be used. Empirical Bayes approaches, also known as maximum marginal likelihood methods [51], represent one approach for setting hyperparameters.

4.2.1 Methodology Comparisons

Within the Bayesian discrete-use system reliability analysis domain, the method proposed by Hall et al. [52] is routinely used to estimate the initial system reliability after the first test phase. To assess Hall's approach's robustness compared to the proposed alternative method, simulation was utilized to generate observed failures and corrective actions within a range of hypothetical systems.

4.2.2 Bayesian Hyperparameters

Within Bayesian statistics, a hyperparameter is a parameter of a prior distribution, with the term hyperparameter being used to distinguish them from the parameters of the model for the underlying system under analysis.

Typically a prior comes from a parametric family of probability distributions partly for explicitness (so one can write down a distribution and choose the form by varying the hyperparameter, rather than trying to produce an arbitrary function), and partly so that one can alter the hyperparameter, particularly in the method of conjugate priors, or for sensitivity analysis.

4.2.3 Conjugate Priors

When using a conjugate prior, the posterior distribution will be from the same family. Still, it will have different hyperparameters, which reflect the added information from the data: in subjective terms, one's beliefs have been updated. For a general prior distribution, this is computationally very involved. The posterior may have an unusual or challenging to describe form. Still, with a conjugate prior, there is generally a simple formula relating the hyperparameters of the posterior to the values of the hyperparameters of the prior. Thus the computation of the posterior distribution is straightforward.

4.2.4 Sensitivity Analysis

A key concern of Bayesian proponents, and criticism by critics, is the posterior distribution dependence on the selected prior. Hyperparameters address this by permitting their easy variation to examine how the posterior distribution (and various statistics, such as credible intervals) vary. A reliability practitioner can see how sensitive their conclusions are to their selected prior assumptions. This process is often called sensitivity analysis.

Similarly, a reliability practitioner may use a prior distribution with a hyperparameter range, perhaps reflecting uncertainty in the correct prior.

4.3 *Hall's Method*

Hall identified and used the likelihood function given by

$$L_k(m, \vec{t}|\vec{P}) \equiv \prod_{i \in obs} [(1 - P_i)^{t_i - 1} \cdot P_i] \cdot \prod_{i \in obs'} (1 - P_j)^T \tag{45}$$

where $\prod_{i \in obs}[(1 - P_i)^{t_i - 1} \cdot P_i]$ is the joint geometric density function of a random sample of size m, which represents the probability that the observed failure modes on trials $\vec{t} \equiv (t_i: i \in obs)$, for example, the term $(1 - P_i)^{t_i - 1} \cdot P_i$ is the geometric probability of observing failure mode i on trial t_i and $\prod_{i \in obs'}(1 - P_j)^T$ is the joint geometric reliability function of a random sample of size $k - m$, representing the probability that the unobserved modes do not occur in T total trials.

Equation (45) represents the likelihood that the m observed failure modes occur with the failure mode first occurrence trials $\vec{t}$ and that the unobserved failure modes do not occur before the end of the test phase, that is, by trial T. Hall interpreted the P_i in Equation (45) as an independent and identically distributed Beta random variable, which leads to the marginal likelihood function

$$\begin{aligned} L_k(n, x) &\equiv E\left[L_k(m, \vec{t}|\vec{P})\right] \\ &= \left[\frac{\Gamma(n) \cdot \Gamma(n - x + T)}{\Gamma(n - x) \cdot \Gamma(n + T)}\right]^{k-m} \cdot \sum_{i=1}^{m} \frac{\Gamma(x + 1) \cdot \Gamma(n - x + t_i - 1)}{\mathrm{B}(x, n - x) \cdot \Gamma(n + t_i)} \end{aligned} \tag{46}$$

Deriving the limiting behavior of the likelihood function results in the initial reliability estimate for a system with an infinite number of failure modes at the commencement of the test phase before any corrective actions as

$$\hat{R}_{\infty,I} = exp\left(\frac{-m}{\hat{n}_{\infty} \cdot [\psi(\hat{n}_{\infty} + T) - \psi(\hat{n}_{\infty})]}\right) \tag{47}$$

where $\hat{n}_{\infty}$ is the estimate of the mode failure probability Beta(n,x) n shape parameter from the observed failure data.

In conjunction with the system-level initial reliability estimate, Hall reduced the estimation procedure to solving only one unknown within a single equation such that

$$\sum_{i=1}^{m}\left(\frac{1}{\hat{n}_{\infty} + t_i - 1}\right) = m \cdot \left[\frac{\psi'^{(\hat{n}_{\infty})} - \psi'^{(\hat{n}_{\infty}+T)}}{\psi(\hat{n}_{\infty} + T) - \psi(\hat{n}_{\infty})}\right] \tag{48}$$

Equation (48) may be used to simply solve Equation (47) when Hall's methodology is adopted in a relatively straightforward fashion.

4.4 *Proposed Alternative Approach*

In the proposed alternative approach, we utilize a different likelihood function for a single failure mode given by

$$L(t|p_i, v_i, d_i, T) \propto$$
$$p_i^{n_{i,1}}(1-p_i)^{v_i - n_{i,1}}(1-p_i)^{(1-d_i)[T-v_i-(n_i-n_{i,1})]}[1-(1-p_i)^{1-d_i}]^{n_i-n_{i,1}} \tag{49}$$

For failure mode probability p_i and a total number of T test demands with arbitrary corrective actions, we assume for the i^{th} failure mode, there are n_i failures on trials $t = (t_{(i,1)}, \dots, t_{i,ni})$ with corrective action implemented on trial v_i and a fix effectiveness factor of d_i. We also further assume there are $n_{i,1}$ mode failures before corrective action implementation and $n_{i,2} = n_i - n_{i,1}$ mode failures after corrective action.

This likelihood function differs from other approaches, including Hall's, in that previously, the fix effectiveness factor has been employed by directly scaling the probability of failure for a failure mode. This proposed approach yields a mode failure reliability after corrective action of

$$R_{i,new} = {R_i}^{(1-d_i)} \tag{50}$$

This approach provides greater consistency between continuous-use and discrete-use reliability growth projection models. Continuous-use models assume a constant failure rate for each failure mode, resulting in an expression of reliability using an exponential distribution. A log-transform of a continuous-use model will result in the failure rate remaining after corrective action. Log-transforming the exponential representation in Equation (50) results in

$$\lambda_{i,new} = (1 - d_i)\, \lambda_i \tag{51}$$

Equation (51) fits the accepted definition for a fix effectiveness factor within continuous-use system reliability growth projection models. It also allows for all failures (both before and after the corrective action) to be used rather than just the first occurrence of a mode. This is a critical delineator between Hall's and the proposed approach. It is also why the fix effectiveness factor is in the proposed likelihood function and not in Hall's model.

In the proposed method, the marginal likelihood for a single failure mode is the denominator of the posterior distribution (the observed model evidence in a Bayesian context) and is given by

$$p(n_i) = \frac{\Gamma[a+b]}{\Gamma[a]\Gamma[b]} \sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j \frac{\Gamma\left[a+n_{i,1}\right]\Gamma\left[b+v_i-n_{i,1}+\tau_{i,j}\right]}{\Gamma\left[a+b+v_i+\tau_{i,j}\right]} \tag{52}$$

where $\tau_{i,j} = (1-d_i)\left(T-v_i-n_{i,2}+j\right)$.

The total system likelihood over k failure modes that exist in the system is then

$$L(n) = \prod_{i=1}^{k} \frac{\Gamma[a+b]}{\Gamma[a]\Gamma[b]} \sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j \frac{\Gamma\left[a+n_{i,1}\right]\Gamma\left[b+v_i-n_{i,1}+\tau_{i,j}\right]}{\Gamma\left[a+b+v_i+\tau_{i,j}\right]} \tag{53}$$

Assuming that m failure modes are observed during the test, we establish an additional mode failure probability prior distribution parameter for the Beta(a,b) parameters such that

$$\tilde{n} = a + b \tag{54}$$

and then let the mean prior reliability of the system be denoted as

$$R_I = \prod_{i=1}^{k} 1 - \frac{a_i}{a_i + b_i} \tag{55}$$

Taking the sum of the log-likelihood terms for the m observed failure modes and the $k - m$ unobserved failure modes, parameterizing in terms of the prior system-level mean R_I and the $\tilde{n}$ parameter, and then taking the limit as k becomes large results in

$$\ell_\infty(n) = \lim_{k \to n} \ell(n) = \sum_{i=1}^{m} \log\left(\log R_I^{-\tilde{n}} \prod_{q=1}^{n_{i,1}-1} (n_{i,1} - q) \sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j \frac{\Gamma[\tilde{n} + v_i - n_{i,1} + \tau_{i,j}]}{\Gamma[\tilde{n} + v_i + \tau_{i,j}]} \right) + \log R_I^{\tilde{n}} \left[\psi(\tilde{n} + T) - \psi(\tilde{n})\right] \tag{56}$$

Taking the derivative of Equation (56) with respect to R_I and $\tilde{n}$ yields the equations for the empirical Bayes estimates

$$\sum_{i=1}^{m} \frac{\sum_{j=0}^{n_{i,2}} \left(\prod_{q=1}^{n_{i,1}} \frac{1}{\tilde{n} + v_i + \tau_{i,j} - q} \right) \sum_{q=1}^{n_{i,1}} \frac{1}{\tilde{n} + v_i + \tau_{i,j} - q}}{\sum_{j=0}^{n_{i,2}} \sum_{q=1}^{n_{i,1}} \frac{1}{\tilde{n} + v_i + \tau_{i,j} - q}} = m \cdot \left[\frac{\psi'^{(\tilde{n})} - \psi'^{(\tilde{n}+T)}}{\psi(\tilde{n} + T) - \psi(\tilde{n})} \right] \tag{57}$$

and

$$R_I = \exp\left[\frac{-m}{\tilde{n}[\psi(\tilde{n} + T) - \psi(\tilde{n})]} \right] \tag{58}$$

Solving Equation (57) for $\tilde{n}$ using the empirical data from the test phase permits the solving of Equation (58) to ascertain the empirical Bayes initial reliability estimate.

4.5 *Performance Comparisons*

For initial testing of the proposed method, various simulations of a range of hypothetical discrete-use systems and their failures against a known true system baseline were conducted. The constructed simulation algorithm, which was examined and coded via the use of the Python programming language:

- sampled from the Beta distributions chosen to represent the hypothetical system mode failure probability and corrective action effectiveness,

- a uniform distribution selected to randomize corrective action implementation times, with all modes subjected to corrective action either during or after the test phase, and

- a Bernoulli distribution was chosen to represent the proportion of modes subjected to corrective action during and after the test.

This approach permits the two methods' performance to be assessed under a range of possible circumstances in a carefully controlled manner. The outcomes of each simulation were then compared against the true initial reliability of the hypothetical system.

4.5.1 Arbitrary Corrective Action Approach

A delayed corrective action regime is one where all corrective actions are postponed until test phase completion using a "Test-Find-Test" approach [19]. An arbitrary corrective action approach is commonly used in developmental testing. It employs a combination of delayed corrective actions and corrective actions applied to some failure modes during the test. The application of corrective actions during a test occurs when a “Test-Fix-Test” methodology is adopted. Practically speaking, immediate corrective action implementation is not usually possible, nor is delaying a test until individual mode corrective actions can be applied. Typically, some corrective actions are applied as soon as practicable while the test phase proceeds.

Table 11 describes the simulation parameters for five hypothetical comparison cases of a discrete-use system tested in a single test phase employing an arbitrary corrective action approach.

Table 11: Arbitrary corrective action approach simulation parameters.

Case	*k*	*T*	*μ*	*var*	*c*	*e*	*P(fix)*	*CL*
1	250	2500	0.001	0.00002	0.7	0.01	0.3	0.8
2	500	2500	0.001	0.00002	0.7	0.01	0.2	0.8
3	100	1000	0.002	0.00005	0.7	0.01	0.2	0.8
4	25	500	0.005	0.00015	0.6	0.01	0.7	0.8
5	25	100	0.003	0.00005	0.8	0.01	0.2	0.8

Within the five cases presented in Table 11, k represents the total number of failure modes in the simulated system, while T is the total number of test phase demands. μ and var are the failure mode probability mean and variance population parameters, respectively, while c and e represent the fix effectiveness factor mean and variance population parameters. CL is the confidence limit utilised in comparing the distributions arising from the simulation, and $p(fix)$ is the probability of corrective action implementation to mitigate observed failure modes.

Table 12 displays the simulation outcomes. In this table, m and n identify the number of observed failure modes and the number of observed failures, respectively. For each of the reliabilities described (true initial reliability, Hall's estimate, and proposed empirical Bayes estimate), LCB and UCB indicate the upper and lower confidence bounds calculated from the confidence limit. μ denotes the approach's mean initial reliability value. The estimator distribution coverage of the true distribution is characterized by Cov, while the absolute relative error between the estimator and the true mean is indicated by Err.

In most simulation cases, the proposed approach resulted in superior coverage of the true reliability distribution compared to Hall's approach. The distribution mean values remained consistently close for all three comparison distributions. The absolute relative error between the means identified by Hall's approach and the true distribution and the proposed approach and the true distribution remained small. The proposed approach tended to produce slightly smaller absolute relative error results when tested across various circumstances.

Table 12: Arbitrary corrective action simulation case results.

			True Initial Reliability				*Hall's Approach*					*Proposed Approach*				
Case	***m***	***n***	***UCB***	***μ***	***LCB***	***Sp***	***UCB***	***μ***	***LCB***	***Cov***	***Err***	***UCB***	***μ***	***LCB***	***Cov***	***Err***
1	44	50	0.85	0.78	0.71	0.14	0.88	0.77	0.66	1.40	0.02	0.82	0.76	0.71	1.04	0.02
2	88	98	0.69	0.61	0.53	0.15	0.73	0.62	0.52	1.12	0.02	0.66	0.59	0.52	1.09	0.03
3	23	31	0.89	0.82	0.74	0.15	0.92	0.79	0.66	1.55	0.04	0.87	0.81	0.74	0.99	0.01
4	9	11	0.95	0.88	0.81	0.14	0.99	0.84	0.70	1.78	0.04	0.94	0.90	0.87	0.58	0.03
5	4	4	0.97	0.93	0.89	0.09	1.00	0.90	0.75	2.56	0.04	0.97	0.94	0.90	0.81	0.01

Overall, the proposed approach offers better coverage of the true distribution with less likelihood of significant errors when point estimates must be used. The proposed approach maintains a similar absolute mean relative error when distribution means are compared.

4.5.2 Testing Assumptions

In developing the proposed approach, several assumptions were made. It is essential to assess the suitability of the proposed empirical Bayes estimates approach to validate these assumptions against the observed failure modes during testing. Failure to do so will undoubtedly result in erroneous estimates of both the prior system-level mean R_I and the $\tilde{n}$ parameter. This will consequently result in errors in the Bayesian-based reliability credible intervals when assessing current system-level reliability and planning for future reliability growth test phases.

In particular, it is essential to assess the goodness of fit of the mode failure probability, which we have assumed to be of the form

$$p(p_i) = \frac{\Gamma(a+b)}{\Gamma(a)+\Gamma(b)} {p_i}^{a-1}(1-p_i)^{b-1} \tag{59}$$

If the model fails to represent the empirical data, significant errors may arise. In this case, alternative approaches may prove more useful. A Bayesian Chi-squared test can be developed, but the visual plot of observed versus prior expected number of modes described in Section 3.14 tends to provide a useful indication of model fit issues.

An estimate of the expected number of failure modes from the defined mode failure probability Beta distribution for an infinite number of modes is defined as

$$\hat{\mu}(t) = m \cdot \left[\frac{\psi(\tilde{n}+t) - \psi(\tilde{n})}{\psi(\tilde{n}+T) - \psi(\tilde{n})}\right] \tag{60}$$

where $\hat{\mu}(t)$ is the expected number of observed modes. Equation (60) conveniently matches that identified by Hall et al. [52] [32]. This management metric provides graphical insight into the goodness of fit of the model. This is achieved by plotting the actual cumulative number of failure modes observed during testing versus trials against the expected cumulative number of failure modes versus trials given by the model.

Figure 24 demonstrates the graphical goodness of fit approach.

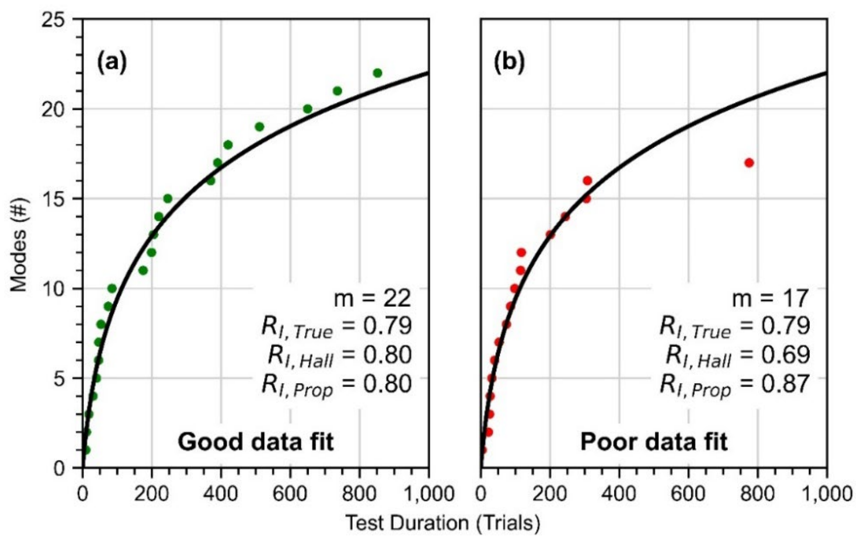


Figure 24. Graphical goodness of fit approach examples.

Figure 24 shows the failure data and the expected number of failure modes plot for two individual simulations of Case 3, detailed in Table 11. Figure 24(a) displays a situation where the model data fits the proposed practical estimation method well. In this case, we expect our empirical initial reliability estimate to represent the unknown system-level true reliability closely with minimal error. This is true for both Hall's and the proposed models. Figure 24(b) illustrates a poor fit between the observed data from a test and the expected values. In this case, the estimate of initial system-level reliability has proven erroneous for both models.

4.6 *Conclusions*

The method proposed in this chapter provides a framework for empirical Bayes estimation of initial discrete-use system-level reliability during reliability growth testing as part of system development. The proposed method provides superior results compared to Hall's approach under various circumstances and equivalent outcomes under others. In particular, the absolute relative error between the estimator and true means is typically less when the proposed method is used.

Chapter 5: Proposed Methodology Modeling Case Study

5.1 *Introduction*

The purpose of this chapter is to introduce a short case study involving the reliability growth planning aspects presented throughout this thesis and their application to a hypothetical discrete-use system. The approaches presented are demonstrated on a complex military system. The methodologies are extended to permit system modeling through developmental, acceptance, and operational testing.

The methods described in Chapter 3 and Chapter 4 are used to develop a prior distribution on the system-level initial reliability. The initial reliability estimator is used to develop test phase reliability growth estimates for a three-phase developmental test.

For acceptance or demonstration testing, it is assumed for this case study that there is no reliability degradation from the developmental reliability growth program. There are no assumed differences between developmental and demonstration testing environments, terrain, weather effects, or representative system use by representative operators. The demonstration test activities will be conducted in an identical location with similar representative operators as developmental testing. Further, it is also assumed that, while no corrective actions will be conducted during demonstration testing, some corrective actions may be implemented between demonstration and operational testing to address any new critical failure modes observed in demonstration testing, but not developmental testing.

A degradation parameter is used to model the degradation in reliability from developmental testing to operational testing. Uncertainty in the degradation factor is modeled using the maximum entropy principles outlined in Chapter 3 to arrive at an approximate Beta(a,b) distribution useful for planning purposes.

5.2 *System Details*

The system in question is a surface-to-air missile. The hypothetical system is shown below in Figure 25.

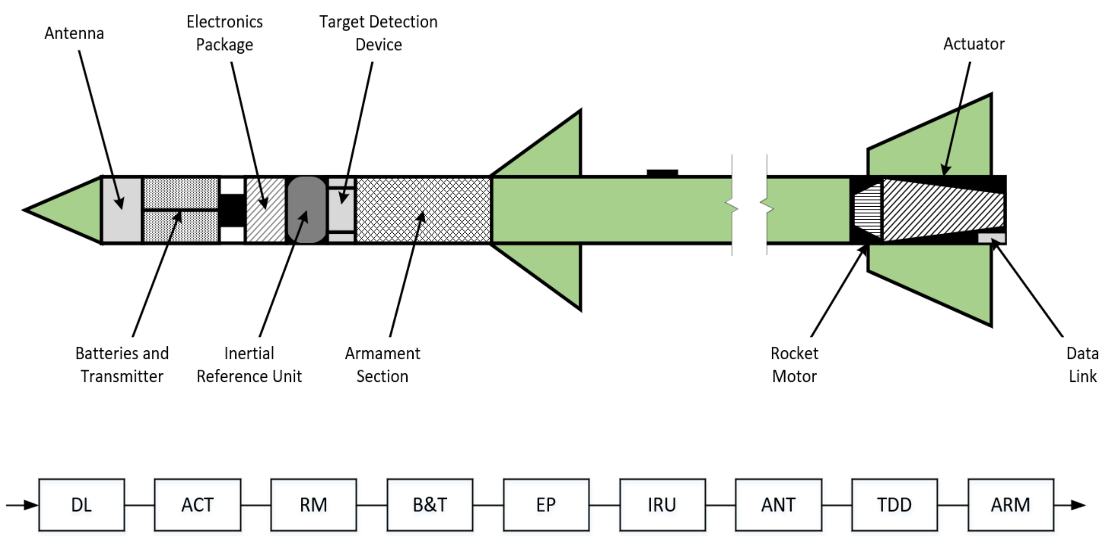


Figure 25. Hypothetical surface-to-air missile system.

The system comprises nine sub-systems that function in series, as indicated by the block diagram within Figure 25. The sub-systems include:

- a datalink (denoted as DL in the block diagram within Figure 25),
- the actuator (identified as ACT),

- the rocket motor (RM),
- several batteries and a transmitter (B&T),
- an onboard electronics package (EP),
- an inertial reference unit (IRU),
- an antenna (ANT),
- the target detection package or device (TDD), and
- the explosive armament section (ARM).

If any of the identified subsystems fail, then the entire system will fail, and the missile will either not launch, fail to acquire the target, fail to maintain target tracking, not maintain stable flight to the target or fail to destroy the target.

5.3 *Modeling Facts and Assumptions*

The following facts are known regarding the proposed developmental, demonstration, and operational test program.

5.3.1 Reliability Developmental Test Schedule and design

Management has directed that a total of 200 hundred systems are allocated to developmental reliability growth testing. No other systems will be made available for developmental testing. The growth program is expected to be completed over nine months. It includes three growth test phases with significant corrective action periods

(CAP) at the end of each phase to ensure sufficient time for design changes and implementation.

During each phase, the following test system demands will be scheduled:

- developmental test phase 1 – 50 systems,
- developmental test phase 2 – 62 systems, and
- developmental test phase 3 – 88 systems.

5.3.2 Reliability Demonstration Test Schedule and Design

A total of 40 systems will be made available for demonstration testing. No other systems will be made available.

Corrective actions may not be applied to observed failures during the reliability demonstration test; however, observed critical failure modes or failure modes not previously observed might be addressed after the test and before the scheduled operational test event.

5.3.3 Operational Test Schedule and Design

Following demonstration testing and the application of any necessary corrective actions, a short duration operational test will be conducted utilizing representative operators in a use representative environment against representative threats.

The operational test activity will have only 12 systems allocated.

5.3.4 Assumptions

The following assumptions have been made to facilitate reliability growth program design:

- Developmental test management strategy (MS_{DT}) – 0.95
- Mean fix effectiveness factor – 0.80
- Probability of corrective action during test – 0.30
- Reliability goal for developmental testing – 0.95
- Post-demonstration test management strategy (MS_{DemoT}) – 0.10

5.4 *System Reliability Growth Program Design*

The purpose of this section is to outline how the proposed methodology may benefit the development of a sound reliability growth program that conveys the range of possible outcomes to a decision maker concerning developmental testing.

5.4.1 Preliminary reliability estimation

Historical testing of similar systems, the manufacturer's previous test data for systems or subsystems, and expert opinion are typically used to estimate preliminary reliability before any testing.

For this case study, this data has been aggregated to provide a preliminary reliability estimate of 0.60 with an expected variance of 0.004. Using the principle of maximum entropy approach described in Chapter 3, the model preliminary reliability estimate distribution may be produced, thereby providing a practical early initial starting point for the planning of the total growth program.

Figure 26 displays the preliminary reliability estimation distribution.

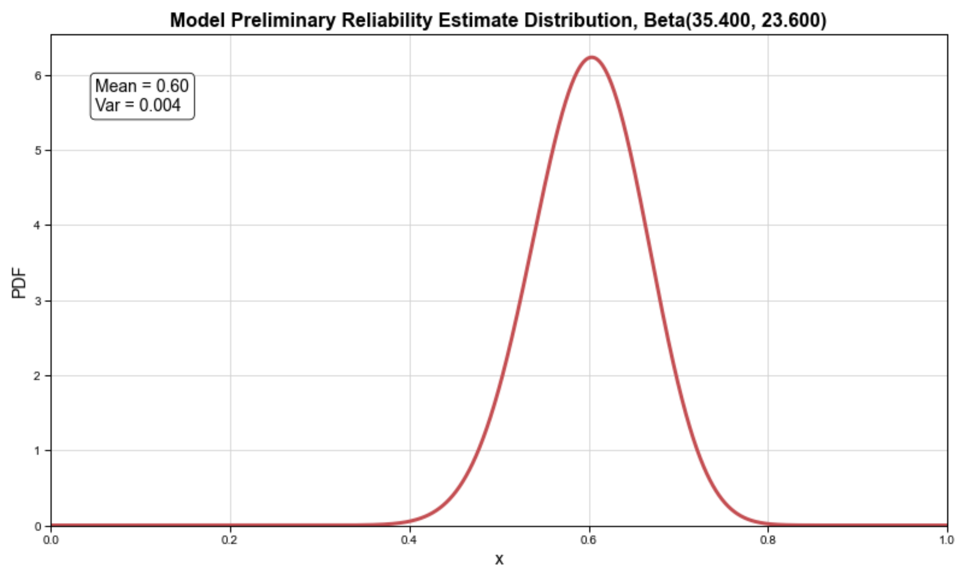


Figure 26. Preliminary reliability estimate distribution.

It is important to note that our belief of the preliminary reliability is untested and used only to establish the expected test entry system reliability for planning purposes. The true initial system reliability is unproven at this early stage.

Also, note that the preliminary reliability estimate is particularly low for an expected high reliability system. However, the low initial reliability chosen helps highlight the clear reliability growth that is expected between test phases for the purposes of this thesis. Choosing a higher preliminary reliability estimate will only demonstrate small reliability gains between phases, making the approach's utility more challenging to convey.

5.4.2 Reliability Growth Testing – Phase 1

Knowing the estimated system-level reliability upon test entry, it is now straightforward to estimate the expected posterior reliability at the conclusion of the first phase of developmental testing.

From our preliminary reliability estimate, we can calculate the mode failure probability Beta(a,b) distribution a and $\tilde{n}$ parameters. Given that T the total number of phase test demands is known, we can now utilize simulation and Equation (21) displayed in Chapter 3 to arrive at our reliability expectation distribution at the conclusion of reliability growth testing phase 1 and after all corrective actions have been implemented.

Figure 27 displays the initial system reliability expectation developed from the model and simulation, given the preliminary reliability estimate. Note that the initial reliability estimate derived from simulation closely reflects the preliminary reliability estimate distribution in Figure 26.

The dashed line represents the mean reliability expectation in the following plots, while the darker shaded areas represent the two-sided 80% confidence limits.

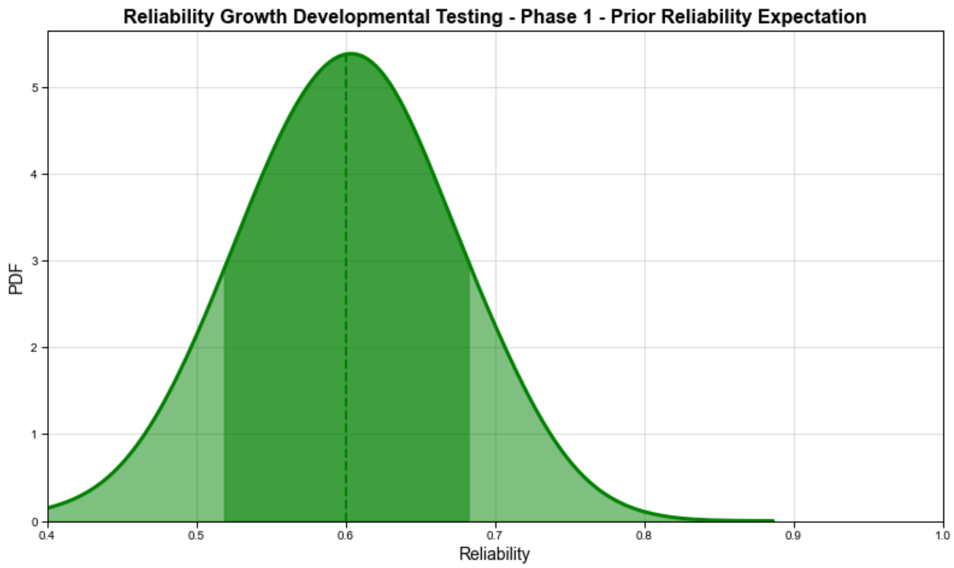


Figure 27. Developmental testing phase 1 initial reliability expectation distribution.

From the simulation, it is identified that, with a two-sided 80% confidence limit, management can expect the initial reliability at test entry prior to phase 1 to be between 0.52 and 0.68. The mean reliability expectation is approximately 0.60.

Adopting a Bayesian learning approach to our planning model, this posterior would now form the prior distribution for phase 2.

Figure 28 displays the reliability expectation posterior distribution at the conclusion of reliability growth testing phase 1. Note that this is after all design changes and corrective actions have been implemented during the post-phase CAP.

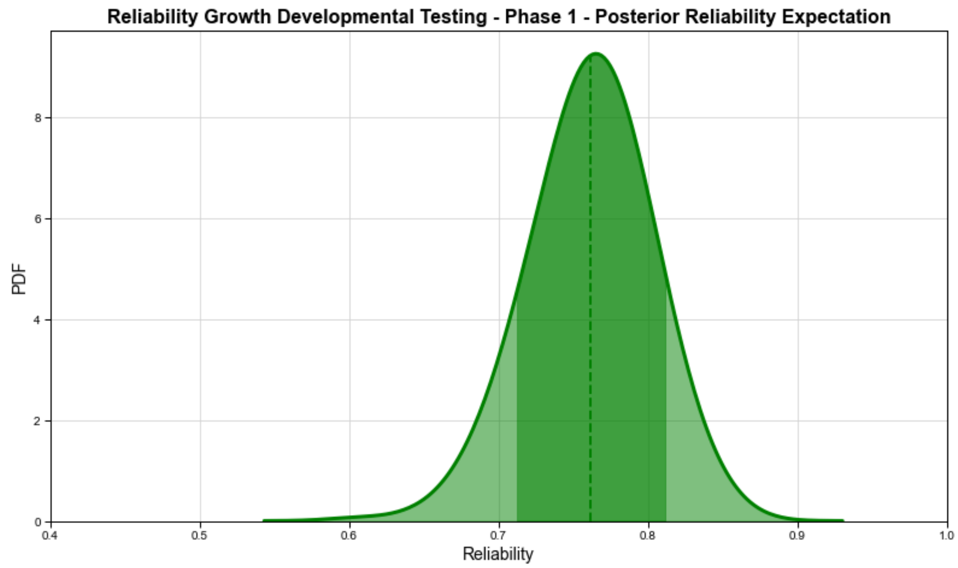


Figure 28. Developmental testing phase 1 posterior reliability expectation distribution.

From the simulation, it is identified that, with a two-sided 80% confidence limit, management can expect the reliability at the completion of phase 1 and after implementation of all corrective actions to be between 0.71 and 0.81. The mean reliability expectation is 0.76.

Additionally, note that the two-sided confidence interval size has decreased from the phase 1 initial reliability expectation to the phase 1 posterior reliability expectation.

5.4.3 Reliability Growth Testing – Phase 2

As stated, the reliability posterior developed during phase 1 simulation now serves as the phase 2 prior distribution.

Figure 29 displays the phase 2 posterior distribution generated through simulation.

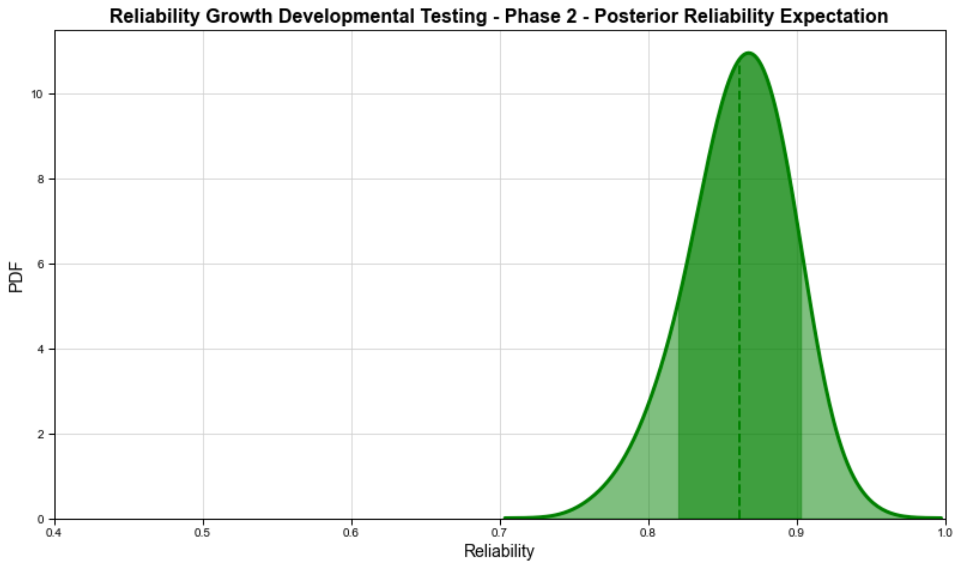


Figure 29. Developmental testing phase 2 posterior reliability expectation distribution.

From the simulation, it is identified that, with a two-sided 80% confidence limit, reliability at the completion of phase 2 and after implementation of all corrective actions should be expected to increase to between 0.82 and 0.90. The mean reliability expectation is 0.86.

Note the further decrease in the posterior reliability distribution between the upper and lower confidence limits. The posterior distribution continues to narrow in terms of spread as more information is utilized in model development. Reliability naturally increases as failure modes are observed, corrective actions are applied, and the probability of mode failure recurrence is reduced.

Model development thus far gives the decision maker (and the reliability practitioner) good visibility of the range of possible test outcomes rather than the limited information offered by a point estimate on a single reliability growth curve.

5.4.4 Reliability Growth Testing – Phase 3

In a similar Bayesian learning manner as previously applied, the posterior distribution developed through simulation for phase 2 now becomes the prior distribution of reliability for reliability growth test phase 3.

Figure 30 displays the phase 3 posterior distribution generated through simulation.

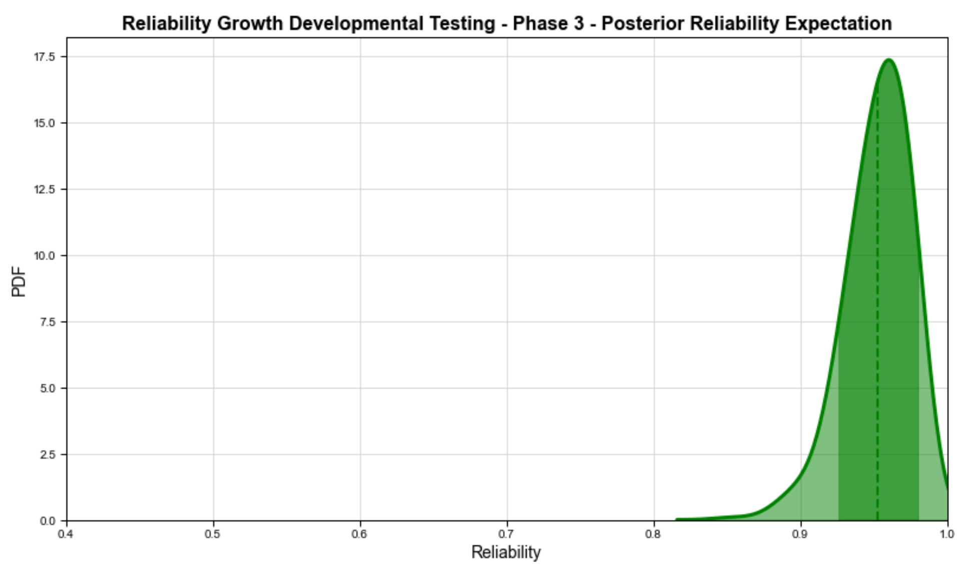


Figure 30. Developmental testing phase 3 posterior reliability expectation distribution.

From the simulation, it is identified that, with a two-sided 80% confidence limit, reliability at the completion of phase 3 and after implementation of all corrective actions should be expected to increase to between 0.92 and 0.98. The final developmental testing mean reliability expectation is 0.95, the specified developmental test reliability goal.

The decision maker responsible for authorizing the developmental test aspects of the reliability program now has complete visibility of the proposed test phases to be conducted, including the range of possible outcomes from the specified management constraints. Should there be some concern that the system will fail to attain the reliability goal, then the decision maker has a range of options available, including:

- increasing the resources committed to assuring reliability growth is achieved,
- increasing the number of test phases together with a corresponding increase in the number of CAPs,
- ensuring the CAPs are of sufficient time length to make certain failure modes can be appropriately investigated and corrective actions developed,
- delaying test entry until a higher initial reliability can be assured by the manufacturer or contractor, and
- altering the goal reliability to maintain realism with available developmental test resources and other constraints (time, schedule conflicts, cost etc.).

5.5 *Demonstration Testing*

Recall that management had constrained the number of systems available for demonstration testing to 40 systems. No other systems will be made available, and this could be due to a number of test constraints, including resources or facilities, availability, schedule, and overall cost.

Additionally, corrective actions may not be applied to observed failures during the reliability demonstration test; however, observed critical failure modes or failure modes not previously observed may be addressed after the demonstration test and before the scheduled operational test event.

Our principle planning assumption for demonstration testing was that, after demonstration testing has concluded, up to 10% of any new or critical observed modes may be addressed via the application of corrective actions.

Similarly to developmental testing bayesian learning, we may use the posterior reliability distribution developed for reliability growth test phase 3 as the demonstration test prior reliability distribution. Figure 31 displays the simulated demonstration test prior reliability distribution after few minor corrective actions were conducted post-developmental but before the commencement of demonstration testing. Note only a slight increase in reliability with the developmental test phase 3 posterior distribution displayed in Figure 30, very similar to Figure 31.

From the simulation results, with a two-sided 80% confidence limit, reliability at the completion of all developmental testing and after implementation of any further outstanding corrective actions should be expected to increase to between 0.93 and 0.98. The final prior demonstration testing mean reliability expectation is 0.96.

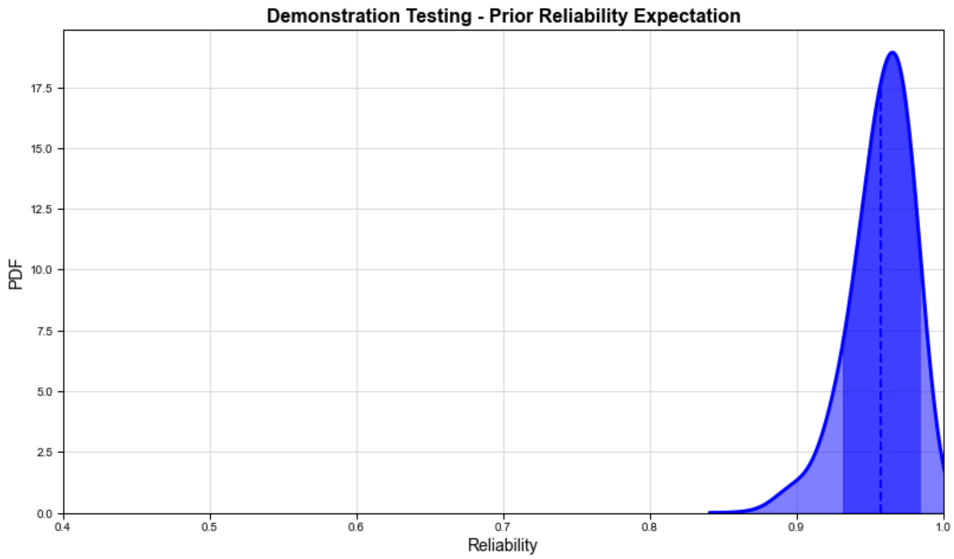


Figure 31. Demonstration testing prior reliability expectation distribution.

Figure 32 displays the demonstration testing posterior reliability expectation distribution. Only slight changes in the distribution may be identified compared to the demonstration testing prior reliability expectation due to only a small amount (no more than 10% of any demonstration testing observed modes) of additional corrective actions post-demonstration testing completion.

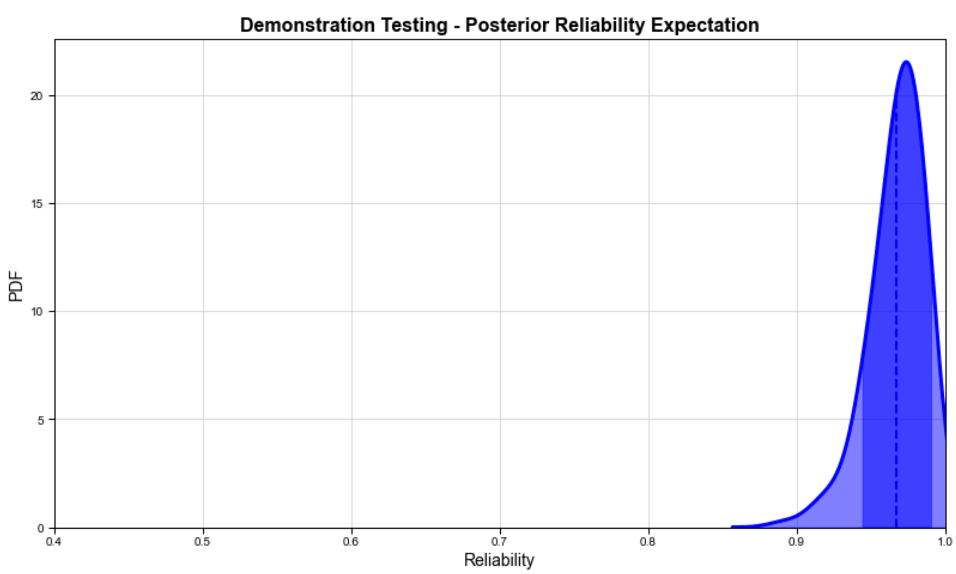


Figure 32. Demonstration testing posterior reliability expectation distribution.

The simulation identifies that reliability at the completion of demonstration testing and after implementation of any resulting minor corrective actions should be between 0.94 and 0.99. The final posterior demonstration testing mean reliability expectation is 0.97.

The data developed during developmental growth and demonstration testing now informs our beliefs on the system reliability moving forward into operational testing.

5.6 *Operational Testing*

Often operational testing is planned and executed with little thought given to previous reliability testing. Operational test authorities also raise concerns when systems fail to demonstrate the reliability proven during acceptance/demonstration testing due to degradation caused by environmental, usage, and system operator changes. The proposed solution here is to utilize the results from the previous demonstration testing by carrying forward this information into operational testing.

5.6.1 Operational Testing Reliability Degradation

Wayne [53] identified a method of combining developmental and operational test data for discrete-use systems that can be employed in a similar manner as the FEF value in Chapter 3. Employing a similar approach to identify the expected reliability in operational test (particularly first-round operational testing where operator use of the system may be subjected to a learning curve response or where operators have limited experience with similar systems) may be useful information for the operational test authority.

Wayne identified the relationship between developmental and operational test data as

$$R_{DT} = {R_{OT}}^{1-\gamma} \tag{61}$$

where R_{DT} is the developmental test reliability derived from the data, R_{OT} is the operational test reliability derived from the data, and γ is the developmental test to operational test degradation factor.

The approach identified can be extended to work in a reverse manner, using a reliability mode degradation factor to estimate the change in reliability when transitioning from developmental and demonstration testing to operational testing.

Like the FEF, it is essential to consider that each failure mode will react differently to the change from a developmental or demonstration test environment to an operational test environment. While some failure modes may only exhibit minimal reliability degradation to the changes in the physical, environmental, and operator-induced test conditions, the reliability of other modes may degrade significantly.

Consequently, the employment of a degradation factor suitable for estimating mode reliability transition between any test environment may be described by

$$R_{\mathrm{FT}_i} = {R_{\mathrm{CT}_i}}^{1-\gamma_i} \tag{62}$$

where R_{CT_i} is the failure mode i reliability observed or estimated from available data for the current test environment, R_{FT_i} is the failure mode i reliability estimated from available data for the future test environment, and γ_i is the modal reliability degradation factor.

The degradation factor may be estimated with the greatest uncertainty by employing a maximum entropy approach similar to that laid out in Chapter 3 Section 3.6.1.

In this case study, the degradation factor is assumed to be a mean value of 0.10 with a variance of 0.001 for each modal reliability. In reality, this may be based on past test history, operational testing degradation results from similar system tests, expert opinion elicitation, or analytical comparisons between the developmental, demonstration, or operational test environments, if such detail were known and available.

Figure 33 displays the case study operational testing reliability degradation estimate derived through the maximum entropy approach.

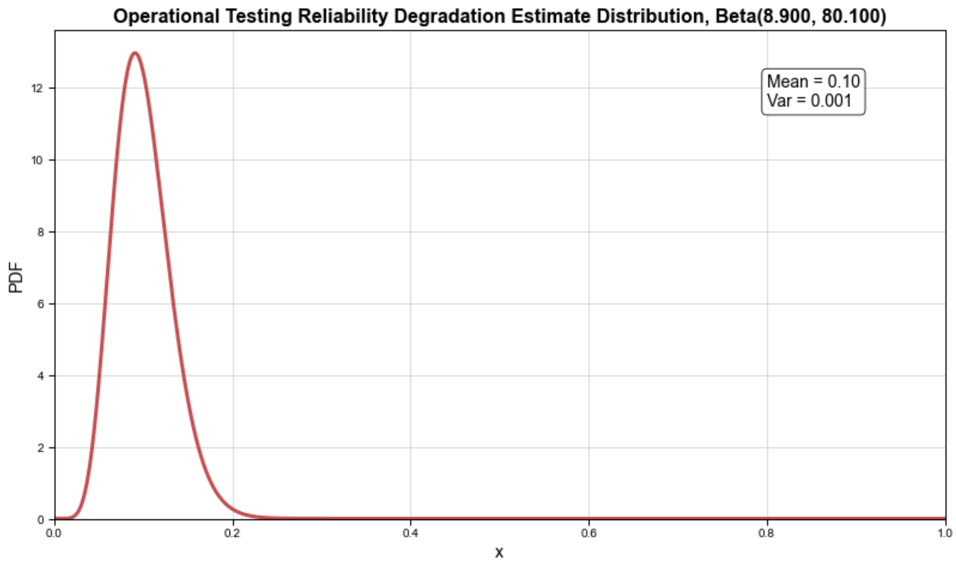


Figure 33. Operational testing reliability degradation estimate.

Entering this data into the simulation through the known Beta distribution parameters (displayed in the plot title in Figure 33) results in the operational test prior reliability expectation distribution shown in Figure 34. This distribution is expected to remain valid for the planned operational testing but may require further analysis resulting in adjustment if the operational testing scenario or environment changes.

The same methodology can now be extended to plan for further operational tests or model system reliability at any stage of the system life cycle relative to changes in the operating context or the use environment.

For capability managers, this projection methodology is powerful and likely to offer significant insights into future system reliability at any time.

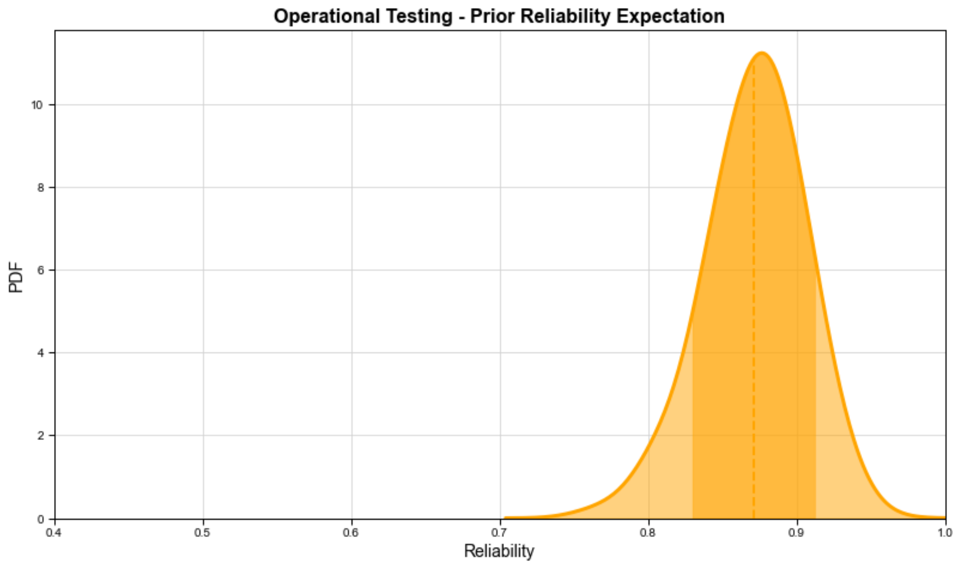


Figure 34. Operational testing prior/posterior reliability expectation distribution.

In this case study, no corrective actions are expected to failure modes observed during operational testing. Consequently, the operational testing prior and posterior reliability distributions are the same.

5.7 *Other Management Metrics*

Based on the results of significant research and simulating hundreds of different outcomes of the proposed Bayesian and simulation approach, several other management metrics may also be considered, including one-sided and two-sided confidence bounded histograms of:

- the number of failures observed during each simulation phase or test event,
- the number of modes observed,
- number of failures observed before implementation of corrective actions,

- number of failures observed after corrective action,
- the expected number of correctable failure modes in total, and
- the number of modes subjected and not subjected to corrective action.

An entire range of management metrics are demonstrated within the Python code detailed in Appendix 2.

5.8 *Discussion*

5.8.1 Methodology Advantages

The proposed reliability growth planning methodology possesses significant advantages over current methods. In particular, the proposed Bayesian approach:

- offers flexibility, and the reliability program can easily be updated as new data and information becomes available simply by adjusting and re-running the simulation;
- offers decision makers insights into the impacts of management decisions on the reliability test program; as the decision maker adjusts their strategies, outcomes can easily be estimated through simulation;
- demonstrates to the decision maker (and the reliability practitioner) the uncertainty surrounding the various elements of the reliability program;

- permits the use and subsequent reuse of all data attained during preceding test events (including activities beyond reliability testing if sufficient data can be captured); and

- displays the range of likely or possible outcomes, based on available information, to stakeholders more precisely than existing models in the literature that portray a simple point estimate at any given time in the reliability growth program.

Reliability communication and risk aspects are covered in greater detail in Chapter 6.

5.8.2 Methodology Disadvantages

The proposed reliability growth planning methodology also exhibits some disadvantages when compared to current methods. Specifically, the proposed Bayesian approach:

- can be computationally intensive depending on program variables and constraints;

- remains predominantly theoretical only in its early stages if insufficient data is available to promote sound conclusions;

- may result in frequent adjustments, particularly in the early stages of model development as new data arrives, or if the initial variable parameters prove unrealistic; and

- may appear to be confusing to decision makers that desire to understand how the reliability projection methodology contributes to the overall plan

5.9 *Conclusions*

The presented case study demonstrates the utility of the proposed approach in developing a methodology suitable for modeling an entire reliability program comprising developmental, demonstration, and operational test phases. Further, the presented case study demonstrates that the approach may also be valid for reliability improvement programs beyond operational testing where optimized system performance is desired.

Chapter 6: Effectively Communicating Developmental System Reliability Growth Plans and Risk[3]

6.1 *Introduction*

Many reliability engineering-related decisions are usually technically advanced, while the decision makers are typically not experts in the field. Corporate leaders, elected government officials, military program managers, project executives, and others in senior positions often sponsor, support, resource, approve and fund reliability growth, tracking, and projection activities. These decision makers are well educated and sincere, and, over the working day, the typical expectation is that they make crucial decisions on various topics in unrelated fields. It is incumbent on reliability practitioners to quickly and efficiently state their case, highlighting any risk. Reliability engineers must learn to present information clearly and succinctly using compelling methods that convince those that are pivotal to success to take action.

However, the global communication landscape is rapidly changing, and effectively communicating reliability engineering information can be derailed by many factors. It is becoming more frequent that engineers face misinformation, purposeful denial, ignorance or indifference, political expedience, and many other problems. In addition,

---

[3] The content of this chapter was presented at the Australian Integrated Project Engineering Congress (IPEC) May 26-28, 2021. Note that the presented paper was tailored to an audience that included those with a limited reliability engineering background. Consequently, some concepts are explained in more detail than would be necessary for a knowledgeable audience. A full-text adapted version of this chapter has been submitted for publishing consideration to the Australian Journal of Multidisciplinary Engineering, a Taylor & Francis Group publication.

the reliability engineering environment has become increasingly complex and reflective of current society. Within the community, diminishing trust in institutions and experts, increasing use of social media as an information source, and the rise of alternative viewpoints inconsistent with evidence are commonplace. The current challenging information landscape can be described as the “post-truth, post-trust, post-expert world” [54]. In a reliability engineering context, this equates to an environment where developmental testing evidence is challenged, trust in professionals is weakened, and the advice of experienced reliability practitioners is held in lower regard.

The purpose of this chapter is to outline the communication approaches that are necessary for motivating decision makers and colleagues in supporting and implementing reliability engineering efforts. This chapter describes how inspiring decision makers to act requires extracting the essence of an argument or concept, crafting key messages, presenting quality risk visuals, bridging knowledge gaps, and creating compelling narratives. The aim is to challenge technical professionals to present information effectively, resulting in superior outcomes.

6.2 *Good Communication Fundamentals*

Most reliability practitioners are adept in a range of “hard” or quantitative skills, those skills they have been trained in and developed as a function of formal education, role, tenure, and position. Hard skills are often quantifiable and require professionals to learn and improve actively. Conversely, the stereotypical engineer is considered deficient in “soft” or professional skills. These interpersonal skills describe how individuals work and interact and are foundational in collaborating and succeeding in the work environment [55]. An essential professional skill for any technical person is communication. Practical communication skills are fundamental to success in many aspects of life. However, communication is not the same as informing [56].

Informing can be considered as the transmission of a message from a sender to a receiver. The message's content typically consists of objective facts, and it is codified independently of the human relationship between the informer and the informed. The message is coded through a system of conventional composition rules, and the sender expects obtainable results.

Communication differs in that a bidirectional sequence of the transmission of messages occurs. Often, parties are considered both senders and receivers. Besides the message being codified through conventional language, the communication actors also send additional messages codified naturally that clarify the relational content between the communication counterparts, such as body language, tone of voice, speech rhythm, and physical posture. Not all messages are transmitted consciously; for example, body language may indicate a different message conveyed during speech alone [57].

6.2.1 Know the Audience

When reliability professionals communicate, the purpose is not what they want to do; instead, in most cases, it is what they want the audience to do as a result of a common understanding. There are many differing world views and perspectives, and different audiences will respond differently to various engineering communication efforts. To communicate effectively and to achieve the purpose, reliability communicators must adapt to their audience. Many communicators assume that a decision maker audience is composed of “people like me.” On the contrary, this is often not the case [58].

Audience segmentation is a method pioneered within marketing used to design and tailor products, services, and messages that satisfy the requirements of targeted groups [59]. Audience segmentation is a crucial activity within an audience analysis that can divide individuals or groups into homogeneous subgroups based upon defined criteria. The process seeks to identify subgroups within a target audience to deliver more tailored messages for more robust connections and is based on a wide range of elements, including predetermined beliefs or values, demographics, psychographics, communication behaviors, and many more [60]. Defining the target audience allows the tailored message to resonate with specific people, resulting in the messenger’s desired outcomes.

6.2.2 Audience Segmentation

Reliability professionals sometimes believe that there is only one decision maker type and a single method of presenting developmental reliability growth plans and risks. However, there are many things reliability practitioners can learn from segmentation studies, and the first is to accept that there is no one decision maker mold.

Kim et al. [61] described four different publics for consideration from a problem/issue perspective. From a reliability communication standpoint, we summarise these into three areas to identify the reliability decision maker types and how they react to information.

Active decision makers often actively evaluate decision alternatives, compute different option values, anticipate various outcomes, and explicitly choose on their own [54]. These are the easiest to communicate with, requiring precise, concise information. The active decision maker will often ask questions, seek clarification and specifically ask about the risk related to various action courses. The relationship between the messenger and the decision maker will usually involve a level of trust in that the reliability practitioner will deliver the best-considered recommendations.

Passive decision makers tend to do what is easy; usually, they decide to either do nothing or follow the recommendations presented to them for decision [54]. The passive decision maker is the most difficult to reach through communication. They do not necessarily understand the transmission, and they may have no interest in the problem or any perceived connection to the outcome. These are the most demanding audiences to deal with as they require efforts to present information in a manner that they do not see as time-wasting. Difficulties arise when importance and criticality are poorly communicated, or the decision maker does not comprehend the value proposition.

Hot-issue decision makers may make decision errors in their quest to address an issue before moving to the next quickly [54]. This audience may rush and accept risks without understanding or considering if a decision is genuinely urgent. They may also tend to rely too heavily on intuition permitting errors by allowing it to outweigh analytical thinking [62]. This audience type requires a carefully crafted message that identifies recommendations, risk, and consequence effects rather than immediate and direct results.

6.2.3 Determine the Strategy

If the communication purpose or audience is unclear, it needs to be clarified as much as possible beforehand. Often when presenting topics to a group, the audience is usually strongly heterogeneous. It includes the “jury” (the person or persons who will decide what action to take), together with “spectators” (those that have no bearing on the outcome). Some situations call for the communicator to primarily address the jury when seeking a result from a single or few decision makers. Others require the communicator to address a larger audience, including the jury. This is typical if broader support for a course of action is required [54].

By understanding the audience, we can develop the strategy. Good communications strategies are based on:

- pretesting the key message with the audience wherever possible and then adjusting,
- shaping the message so that it is relevant to the jury, spectators or both,
- ensuring the message is accessible and understandable in format and language,
- using graphics and images for illustrating technical points,

- confirmation of the received message and understanding, and
- concentrating outcomes as they relate to the audience (for example, answering the question “So what?”).

Distilling the message to its simplest form related to the selected audience is often the most potent way of communicating a complex message. Additionally, simplifying the message ensures that the appropriate strategy is selected that meets the audience’s needs.

6.2.4 Simplify the Message

Having a simple message can often be difficult for engineers and technical professionals. For some who have spent considerable time attempting to provide context, more data, or more precision, it can be challenging to abandon that for a simple cut-through message that can influence a decision maker. A general rule of thumb is to have three but no more than five key messages in any single communication effort [54].

Some communication experts advocate using a ‘message box’ [63], a simplified template for crafting a key message. One such template is the COMPASS Message Box [64], which looks at the problems, benefits, the so what, and solutions to an issue. The Message Box is designed to help take information and communicate it to resonate with the chosen audience. Figure 35 graphically demonstrates the Message Box concept.

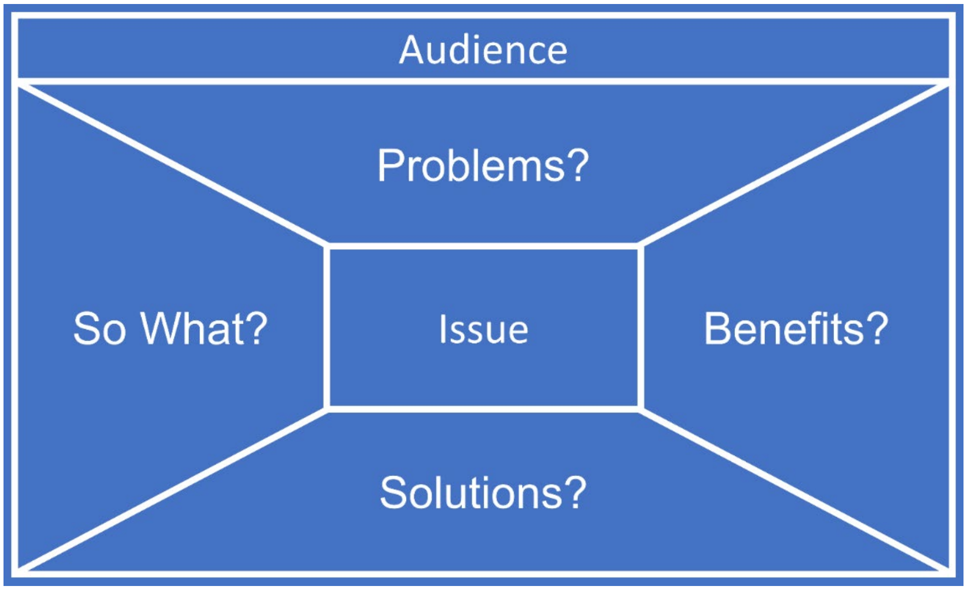


Figure 35. The COMPASS Message Box concept [64].

The message box approach can assist in preparing for interviews or briefs with executives, plan a presentation, outline test, and evaluation outcomes, prepare funding proposals, explain the nature of the work proposed or conducted, and why it matters. Firstly, the communicator identifies the Audience. Then they may move through the five sections of the message box in any order. First drafts are often poorly structured and usually provide too much detail. Outcomes are refined and distilled with each new version. The communicator makes choices about message point importance through each iteration and makes language adjustments until it succinctly captures the key messages that need conveying. By the end of the process, only the essential key informational points should remain for each section.

The elements of the message box include:

Issue: The Issue section identifies the overarching topic that needs addressing. The Issue should be concise and clear, no more than a couple of words or a short phrase, and, as a guide, should approximate something entered into a search engine.

Problem: The Problem is the broader Issue that requires addressing through management action or decision. It typically relates to the communicator's knowledge and area of expertise. The Problem section describes adverse observations and what needs managing, leading into the So What section.

Benefits. In the Benefits section, the communicator lists the positives of addressing the Problem and the expected outcomes if the Solution is implemented. This ties into the So What of why the Audience cares but focuses on taking action.

So What. The So What section is the most critical question that the Message Box helps communicators answer. It describes the impact of the Issue on something the Audience cares about profoundly. The answer to the question changes from audience to audience and requires adjustment accordingly.

Solutions. The Solutions section describes the options to address the Problem. These must be solutions that the Audience can influence or implement. The communicator must revisit why they must communicate with the Audience and their expected outcomes.

An example Message Box for a reliability engineer presenting a project system reliability challenge to a project executive is demonstrated in Figure 36.

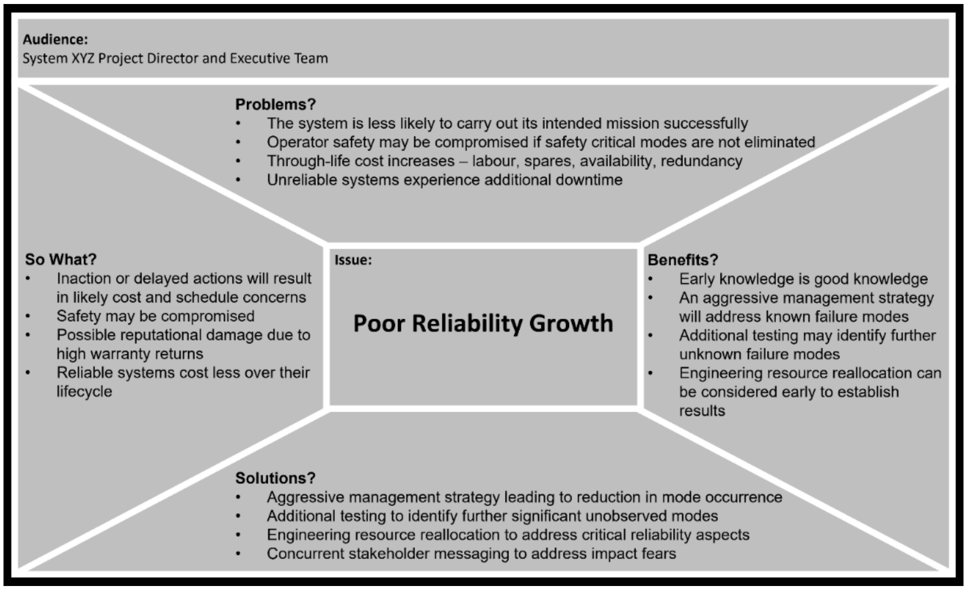


Figure 36. Completed Message Box example.

While relatively simple in its approach, the Message Box concept can be fundamental in determining the communication strategy and crafting the key messages.

6.2.5 Storytelling and Engagement

Storytelling is a two-way interaction between someone telling a story and the listeners. It is a well-known and influential means of communicating messages and engaging audiences [65].

Using storytelling to explain complex concepts has, in the past, not been considered a rigorous method of communicating engineering matters. However, an increasing number of studies show that narratives can help develop trust with an audience and increase knowledge retention and the ability and willingness of audiences to learn and take action. Being easily digested by the human brain, stories help bridge our logos and pathos; when an audience becomes emotionally receptive to facts, chances increase that they will respond and act on the knowledge [54] [65].

Traditionally, engineering knowledge communication focuses on isolated logical ideas with limited context given to the target audience. However, presenting isolated content poses the risk that the audience, particularly the non-expert one, might make inaccurate assumptions when trying to understand the new information [66].

Communicating evidence in an understandable and practically relevant way for stakeholders, for instance, by embedding knowledge in a narrative storyline, has shown to increase an audience's engagement, willingness to act upon the knowledge and then use the evidence as a basis for their decisions [66] [67] [68]. In addition, by placing knowledge into context, stories are more comfortable to process and generate more attention and engagement than traditional logical-based engineering communication [69].

6.2.6 Communication Goals

When reliability practitioners try to help others think the way they do, they need to be precise. Reliability engineering can be a complicated endeavor, but the descriptions and explanations do not need to be. Nevertheless, since reliability professionals are not salesmen, they seek to make decision makers believe what they say because they understand. For reliability engineers, this is the challenge.

Learning how to communicate the essential material without reducing it to the lowest level is critical. Unfortunately, engineers are rarely taught this during their studies, where providing more and more detail is usually seen as the path to success. An essential part of this process is vital message identification –the most critical idea the decision maker must comprehend to initiate action.

6.2.7 Framing

Casting information in a particular light to influence what people think, believe or do is known as framing. Influence may be considered a contentious word in this context, and framing is considered "slanting the facts" or manipulating the truth. Some hold a belief that the only unbiased way to reach an audience is by stating the truth. However, studies have shown that factual information is no better at influencing an audience than information that has no basis [70]. What tips the scales is how facts are presented rather than the actual content.

6.3 *Communicating Reliability Growth Plan Risk*

Targeting a specific communication technique to a particular audience can be challenging. The following section describes how communication may be specifically adapted to address both the reliability practitioner and the audience's requirements to ensure that the message is understood.

The goals of reliability growth planning are to optimize testing resources, quantify potential risks, and permit reliability objectives. The growth plan is an essential management tool in scoping the resources necessary to enhance system reliability and improve the likelihood of demonstrating the system reliability requirement. Critical aspects underlying this process include addressing program schedules, amount of testing, resources available, and the test program's likelihood of achieving reliability requirements.

Figure 37 displays four reliability growth planning models. For comparison purposes, each image represents the same single-phase reliability growth test developed for the same continuous-use system.

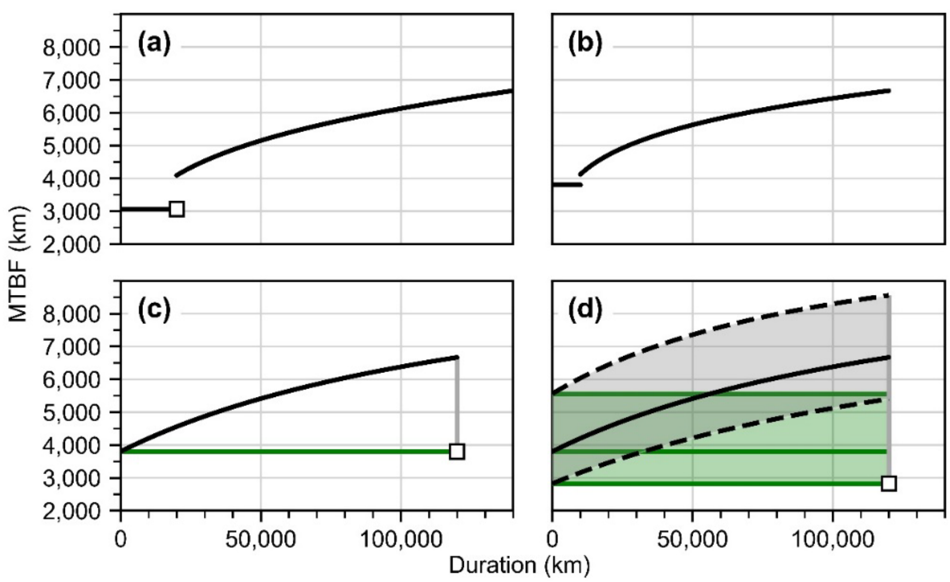


Figure 37. Comparison of reliability growth planning model approaches.

Figure 37(a) displays an example AMSAA Crow Planning Model plan [71]. This model considers the prior estimate of reliability identified in previous activities or test events. This is represented by the horizontal line at 3,064 km MTBF between test commencement to 20,000 duration. At the end of this initial pretest activity, the square marker represents a corrective action period (CAP) between the two activities to address observed failure modes. The jump in reliability as a result of corrective actions can be observed between the two plots. The growth plan curve then extends to the reliability goal.

Figure 37(b) shows the equivalent System Level Planning Model (SPLAN) plan [71] and is somewhat similar to Figure 37(a). Both Figure 37(a) and Figure 37(b), although simplistic, present information that is akin to a point estimate of achieved reliability at any given time during a reliability growth test activity. Neither plot demonstrates the risk associated with the plan or the variance attributed to unknowns such as the initial reliability at test entry, the effectiveness of developed corrective actions to observed failure modes, or the management strategy (MS) in dealing with these observed failures. This presents a challenge to the reliability engineering communicator as a non-technical audience could receive the message that the expected system reliability "will be exactly this value at this given time during the test phase" when this is unlikely to be true. In addition, the audience may struggle to understand that there are unknowns with the proposed test design as the test plan only demonstrates a single path to an expected outcome.

Figure 37(c) displays the equivalent Planning Model Based on Projection Methodology – Continuous (PM2-C) test plan [71]. PM2 curves primarily consist of two components – an idealized curve and MTBF targets for each test phase. The black line represents the idealized curve. This curve is similar to the previously described curves in Figure 37(a) and Figure 37(b). It depicts reliability growth if corrective actions to observed failure modes were applied instantly. The green horizontal line represents the system's stable reliability if all corrective actions are delayed until the end of the current test phase. The grey vertical line represents the growth in system reliability after all corrective actions are applied during the CAP at the end of the test phase. This provides some indication of risk for a decision maker. In reality, we expect system planned reliability to be between the green and black plots as the period between failure observation and corrective action is typically delayed rather than instantaneous.

Figure 37(d) demonstrates the Wayne Improved PM2-C Model test plan [53]. This method improves upon previous models and provides quantification of the inherent uncertainties that exist in reliability growth planning: the initial system reliability, which is unknown at test phase commencement; the MS, a term used to describe the proportion of observed failure modes to be addressed by corrective action; and the fix effectiveness factor (FEF), the percentage reduction in individual mode failure rate achieved by the implementation of corrective actions. This plot depicts upper and lower confidence limits on the expected idealized and delayed corrective action plots, together with their expected values. Consequently, this modeling approach presents significantly more detail to the decision maker in a graphical form that is relatively simple to comprehend, leading to a greater understanding of the nature of the reliability growth program risks.

6.4 *Elements of Plan Design Risk*

When conveying risk to a decision maker, it is essential to consider the plan design risk rather than solely the risk associated with plan execution. Guidelines exist to assist reliability practitioners in formulating an acceptable, risk-considered reliability growth plan.

Table 13 demonstrates an example planning guidelines risk matrix. The employment of risk matrices within the reliability domain may help promote reliability plan risk discussion, providing consistency in prioritizing risks and focusing decision makers on the highest priority risks. However, like any set of guidelines, they require tailoring to meet organizational and contextual environments.

Table 13: Example reliability growth program risk assessment matrix [72].

Category	Low Risk	Medium Risk	High Risk
MTBF Goal (DT)	< 70% of growth potential	70 - 80% of growth potential	> 80% of growth potential
IOT&E Producer Risk	≤ 20%	21 - 29%	≥ 30%
IOT&E Consumer Risk	≤ 20%	21 - 29%	≥ 30%
Management Strategy (MS)	≤ 90%	91 - 95%	≥ 96%
Fix Effectiveness Factor (FEF)	≤ 70%	71 - 79%	≥ 80%
MTBF Goal (DT) / MTBF Initial	< 2	2 - 3	> 3
Time to incorporate and validate fixes in IOT&E units before the test	Adequate time and resources to have fixes implemented and verified with testing or robust engineering analysis	Time and resources for almost all fixes to be implemented and most verified with testing or robust engineering analysis	Many fixes not in place by IOT&E and limited fix verification
Corrective Action Periods (CAPs)	≥ 5 CAPs which contain adequate calendar time to implement fixes	3 - 4 CAPs, but some may not provide sufficient calendar time to implement all fixes	1 - 2 CAPs of limited duration
Reliability increases after CAPs	Moderate reliability increases after each CAP resulting in lower risk curve that meets goals	Some CAPs show jumps in reliability that may not be realised because of program constraints	Majority of reliability growth ties to few considerable jumps in the reliability growth curve
Percent of initial mode failure intensity surfaced	Growth appears reasonable (a small number of the problem modes surfaced do not constitute a large fraction of the initial problem mode failure intensity)	Growth appears somewhat inflated (a small number of the problem modes surfaced constitute a moderately large fraction of the initial problem mode failure intensity)	Growth appears artificially high (a small number of the problem modes surfaced constitute a large fraction of the initial problem mode failure intensity)

Combining organizational or industry-specific guidelines with a risk matrix simplifies the communication focus significantly. Figure 38 displays an example “gauge chart” that harnesses the power of both approaches. The single snapshot that this approach provides significantly simplifies conveying reliability program design risk.

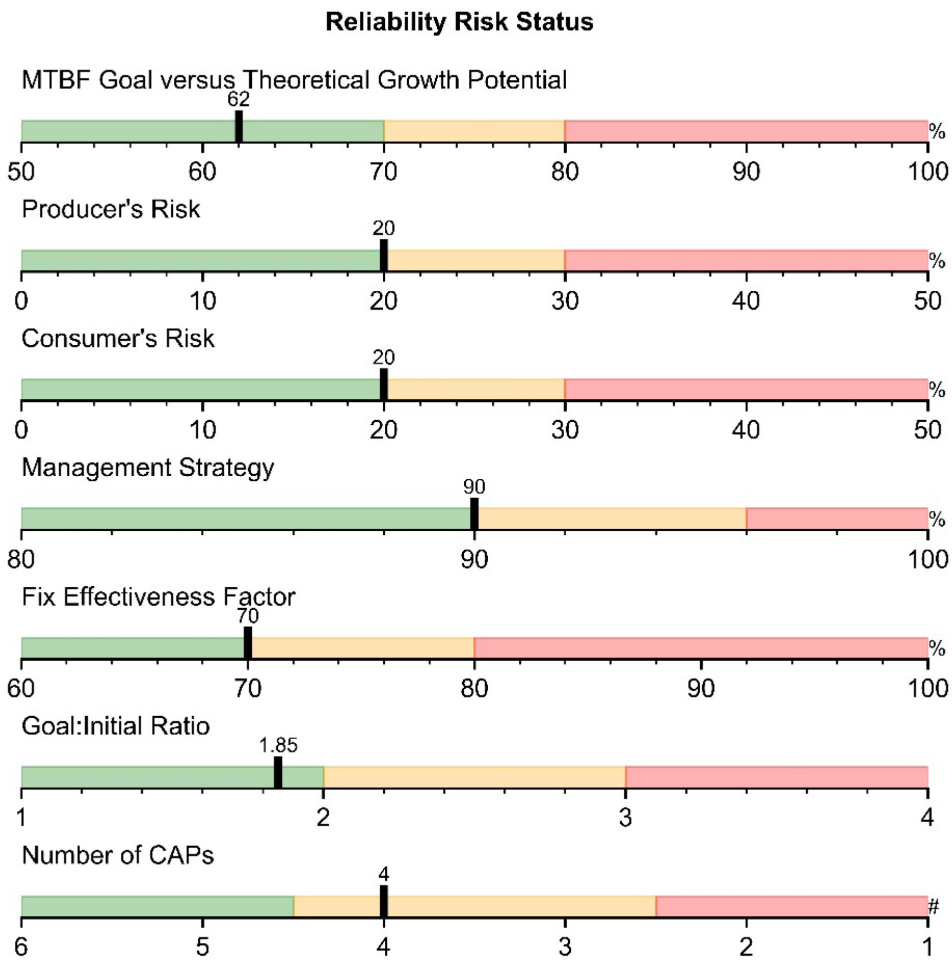


Figure 38. Reliability risk status ‘gauge chart’ example.

6.5 *Conveying the Reliability Growth Plan*

Figure 39 demonstrates an example reliability growth plan for a hypothetical system developed utilizing the Wayne Improved Bayesian PM2-C methodology [53]. The plan is broken into four distinct phases. The decision maker can identify three reliability growth test phases (denoted by RGT) and a single limited user trial (LUT). The decision maker can determine the idealized growth curve and associated upper and lower confidence bounds with a supporting explanation. The phase confidence bounds on the expected reliability if all corrective actions were delayed until the end of the current test phase are identified by the colored phase intervals. The inclusion of CAPs supports the implementation of corrective actions to address observed failures. The initial reliability confidence bounds and expectation value upon test entry are indicated on the y-axis. At the same time, the critical plan variables (FEF, MS and confidence limit, CL) used in its design are annotated in the lower right-hand corner of the plot. The upper dashed plotline annotates the theoretical reliability growth potential (RGP) that results directly from the selected MS and FEF variables. Finally, the goal reliability and the related upper and lower confidence bounds are annotated at the ends of the idealized curve and the confidence intervals.

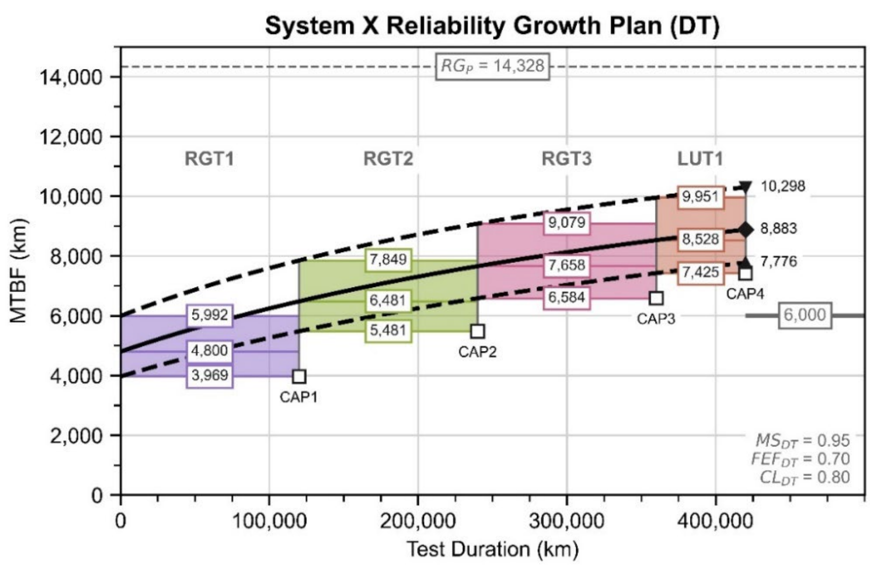


Figure 39. Hypothetical continuous-use system reliability growth plan structure.

The plan represented in Figure 39 can be clearly communicated to a decision maker for consideration. In effect, when seeking decision maker endorsement of the plan and any associated supporting actions, this single figure becomes the proposal when adopting a minimalist approach to reliability communication and attempting to convey the message concisely. The included confidence intervals serve to communicate the risk associated with plan execution.

Conversely, Figure 40 presents the same plan, developed using the standard PM2-C methodology. This plot fails to convey any reliability confidence intervals associated with the plan design, even though both methods present the same underlying information. Simply, this reliability growth plan only presents the reliability expectation value at any point during the proposed test.

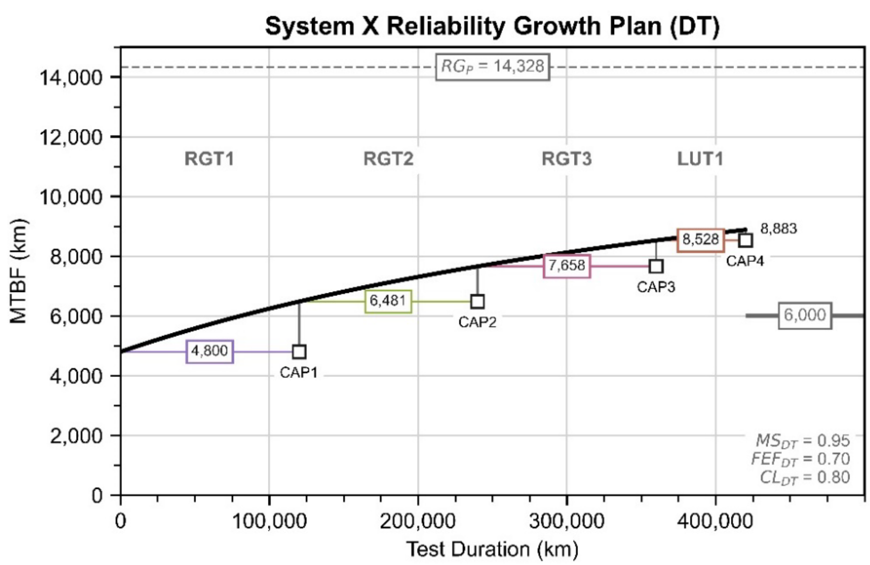


Figure 40. Alternative hypothetical system reliability growth plan structure.

6.6 <u>Communicating Plan Execution Outcomes</u>

During the reliability growth plan execution, the decision maker will usually require periodic progress updates. The updates provide the reliability practitioner with an opportunity to communicate realized or emerging risks. Figure 41 demonstrates an example of how a detailed reliability growth tracking model can be overlaid on the approved reliability growth plan, thereby providing surety to the decision maker that the plan execution is progressing as expected or otherwise. Consequently, the reliability practitioner can propose plan adjustment recommendations that meet evolving decision-making needs with this new information.

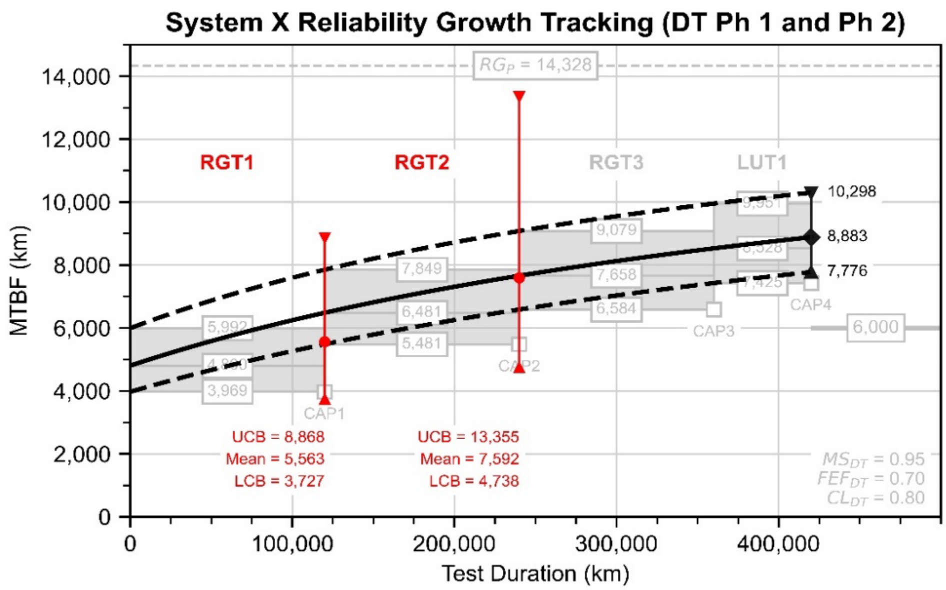


Figure 41. Hypothetical continuous-use system reliability growth tracking.

6.7 *Identifying and Conveying Risk-Based Plan Adjustment Opportunities*

Reliability plan management metrics provide a further opportunity for the practitioner to communicate threats and opportunities to the decision maker for consideration and action. Like the reliability growth tracking graphic illustrated in Figure 41, it is possible to map plan execution progress against the initial plan management metrics to enable the decision maker to optimize resource allocation. Figure 42 displays example plots of four important reliability growth management metrics that may be annotated during plan execution.

The example plots within Figure 42 include the number of B-modes (those failure modes that will be addressed by corrective action) observed during testing so far (Figure 42(a)); the observed mode failure rate (Figure 42(b)); the expected fraction of B-modes not yet observed during testing (Figure 42(c)); and the initial failure intensity attributed to B-modes already observed in testing (Figure 42(d)).

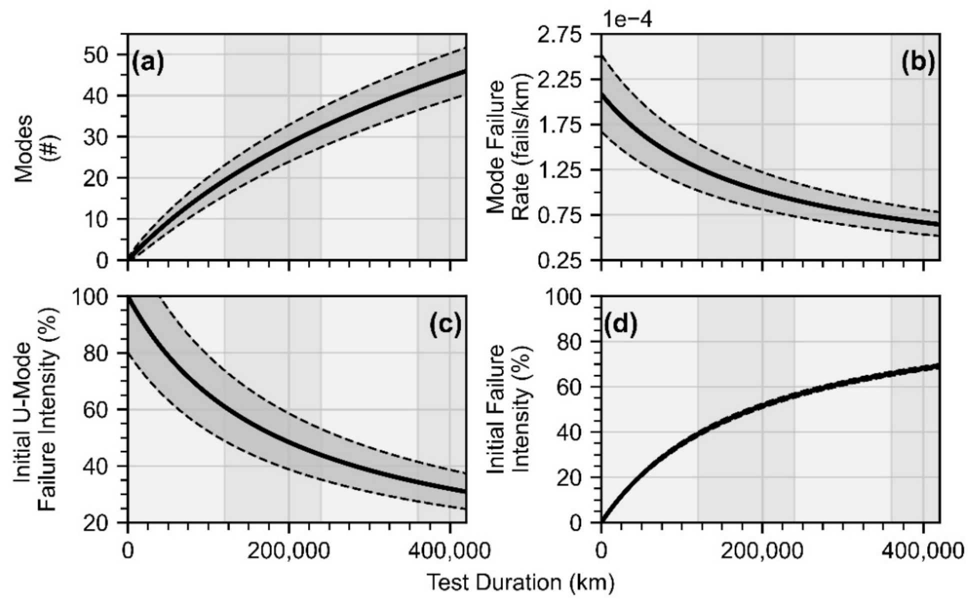


Figure 42. Example reliability growth progress management metrics.

6.8 *Robustness of Reliability Plan Estimates*

The reliability growth plan can only be as good as the estimates that are used in plan design. The purpose of reliability estimation is to determine whether an item has met specific reliability requirements, typically with a stated statistical confidence level.

An estimator of a population reliability parameter is an approximation depending solely on sample information. There are several classes of estimator, which include point, interval, and distribution estimates. A point estimate represents a single estimate of reliability, whereas interval and distribution estimates represent a range of potential reliability values. Thus, interval estimates provide much more information to a decision maker and are preferred when making inferences. Similarly, distribution estimates are derived from a 'distribution function' and convey the most information about a population of reliability test observations. Each has its benefits in communicating reliability and risk as well as shortcomings.

6.8.1 Point Estimates

Decision makers will naturally tend to sample and population observations based on point estimates from observed failure modes, often without knowing it because they are intuitive and easy to understand. The downside of point estimates is that they are often unreliable [73]. Even with a reasonable point estimate, there is very likely to be some error. Other issues also arise. Point estimates provide no information about their sampling distributions, whether the estimator is unbiased or the range of their variances.

The two most common criteria for assessing the reliability of a point estimator are (a) its accuracy or amount of bias and (b) its precision, variation, or concentration [74].

6.8.1.1 Measures of bias

Reliability practitioners use two measures of bias: mean bias and median bias, of which the former is the most popular.

For n samples of a population statistic Θ, the mean estimate, $\overline{\Theta} = \sum \widehat{\Theta} / n$, would be subject to sample variation. However, if there are an infinite number of estimates (which is not possible in a reliability growth context), then the average estimate is of no use; in this case, the bias is expressed as an expected value described as $\overline{\Theta} = E\left[\widehat{\Theta}\right]$.

The mean bias of a point estimator is the expected (mean) difference between the reliability estimate and the true reliability, or $E\left[\widehat{\Theta} - \Theta\right]$. If we denote M as the median of n samples, then the median bias is simply $M\left[\widehat{\Theta} - \Theta\right]$.

For mean symmetrically distributed estimates $E\left[\widehat{\Theta}\right] = M\left[\widehat{\Theta}\right]$ provided that the number of samples is infinite. If the number of n samples is few as it typically is with reliability growth testing, then the possibility that the mean reliability estimator $\overline{\Theta}$ does not approximate the true reliability Θ is high. Hence, we assume that the likelihood of the estimator being significantly biased is also high. Therefore, we consider reliability point estimators as typically unreliable.

6.8.1.2 Measures of concentration

Whilst an ordinary reliability sample mean might be an unbiased estimate of its true reliability mean, this does not imply it is the best estimator of that parameter. For inference purposes, the least variable and efficient estimator provides the most power to discredit a null hypothesis [75].

Suppose reliability testing failure observations have a symmetric distribution because the true median and mean are the same. In that case, the sample reliability median is also an unbiased estimator of the true mean. However, sample reliability medians have a more significant standard error than sample means. A sample mean is the most efficient estimator of the population mean – provided the errors are normal [76].

If testing failure observations are not normally distributed, although the mean is still an unbiased estimator of its parametric value, it is no longer the most efficient. Therefore, depending upon the error distribution, various statistics may provide a less variable estimate of the population mean – bearing in mind their estimates may be biased also [76].

6.8.1.3 Point estimator robustness

Where assumptions are fully met, the arithmetic mean, $\overline{\Theta} = \sum \widehat{\Theta} / n$, is the best possible estimator of its parameter by almost every criterion selected. Unfortunately, this rarely occurs under real-world testing conditions. It is known that, for any reasonable sample size, even relatively minor departures from some of these assumptions, such as slight skew or kurtosis, often make the mean reliability point estimator anything but a reliable estimator of true reliability. However, many reliability practitioners and engineers still maintain its use as a reasonable representation of true reliability.

6.8.2 Interval Estimates

An interval estimate, however, is typically a much more accurate representation of reality. There is some uncertainty when estimating any parameter, which the level of confidence can assess. In contrast to point estimation, which gives a single value, interval estimation uses sample data to estimate an interval of plausible values of a parameter of interest [77].

Interval estimates, while sharing many of the problems of point estimates, tend to be assessed differently. To understand the reasoning and shortcomings of these methods, we must consider how these intervals, and their estimates, are defined.

In essence, a confidence interval $\hat{I}$ estimates a range I which encloses a true reliability statistic Θ. The width of I is set according to what proportion α of all estimates of Θ are excluded from that range (for example, $\alpha = 5\%$). Provided $\widehat{\Theta}$ is distributed symmetrically, I is located centrally about Θ.

This arrangement has two essential properties. We would expect the interval I to enclose the most likely $(1 - \alpha)$ estimate of Θ. Conversely, any estimate of Θ outside that range should be rejected by a comparable test (for example, p-value $p < \alpha$). It then follows that, when $\widehat{\Theta} = \Theta$ and $\hat{I} = I$ the same ought to be true – even if $\widehat{\Theta}$ is distributed asymmetrically. In addition, we could assume that a reasonable estimate of I would perform best.

6.8.2.1 Coverage error

Assuming the confidence intervals are reasonable estimates of I, when these estimated intervals $\hat{\mathrm{I}}$ are attached to estimates of Θ, a predetermined proportion $(1-\alpha)$ (or, for example, 95%) of these intervals are expected to enclose Θ – at least on average. If, as predicted by this model, exactly $(1-\alpha)$ (or 95%) of confidence intervals enclose Θ, this is described as perfect coverage.

The most popular measure of the quality of an interval estimator, known as the coverage error, is simply the difference between the observed and expected coverage [78]. Confusingly, for reasons of mathematical convenience, the formulae for this generally assume that $\widehat{\Theta}$ is distributed symmetrically, and we are calculating the (equivalent 2-tailed) interval between two equal 1-tailed confidence limits. In other words, coverage error assumes a different definition of confidence limits from the one above.

The problem with this measure is that it wholly ignores the length of confidence intervals, or what happens where $\widehat{\Theta}$ is not distributed symmetrically about Θ. Interest in alternate measures of interval estimates and alternate ways of constructing confidence limits is comparatively recent [79] [80] [81].

6.8.3 Distribution Estimates

In principle, a 'distribution function' is a statistic that conveys the most information about a population of test observations or a population of test summary statistics. In simple terms, it describes the probability of observing a particular value of Y that is equal to x ($P[Y = x]$) for each value of x. This is known as the probability mass function (PMF) for discrete distributions, or, for continuous functions, it is the probability density function (PDF) at point x. Alternatively, it may also be considered the probability of observing $\widehat{\Theta}$ within the interval x_1 to x_2 (or $P\left[x_1 < \widehat{\Theta} < x_2\right]$). Among statisticians, the 'distribution function' also refers to the cumulative distribution function of the test statistic($P\left[-\infty < \widehat{\Theta} < x\right]$ or $P[Y = x]$), for each value of x.

The critical consideration for reliability practitioners seeking to communicate reliability test outcomes to decision makers is that formulae are only available to describe a minimal set of theoretical, known distribution functions – which is why they are usually just approximations [82].

A few statistics, those that behave like sums, have distributions that approach normal when calculated from large (or exceedingly large) test samples. When calculated from anything other than large samples, many commonly used estimators have surprisingly complex distribution functions – even when the observations represent a known population distribution, for example, the normal distribution.

When calculated from actual data, the exact distribution of any statistic is highly complex and typically impossible to cope with analytically. For example, likelihood statistics are often assumed to be asymptotically normal, so the distribution of their ratios is tested against a Chi-square distribution. These limit distributions, valid where n approaches infinity, are frequently used as 'parametric' approximations for relatively moderate samples – even though statistical functions approach their asymptotic behavior at widely differing rates.

6.8.3.1 Approximation error

Using some arbitrary but convenient theoretical distribution to approximate the actual distribution of reliability estimates introduces what is known as an approximation error. For instance, approximation errors arise when a statistic is tested, and we assume it is normally distributed when it is actually not.

Although the small-sample distributions of many estimators are complex, we noted above that they often converge asymptotically to known distributions, particularly the normal distribution. Rescaling-transformations and 'studentizing' can reduce, but do not eliminate, approximation errors.

Based on the proceeding discussion on reliability estimates, point estimators are routinely unreliable. They often demonstrate significant bias making them largely unsuitable for communicating reliability statistical information developed during test planning or garnered via test activities. Interval and distribution estimates both have their challenges. While interval and distribution estimates are prone to error, interval estimates typically provide reliability communicators with a suitable helpful tool in conveying reliability growth plan or testing outcome risk to decision makers. In addition, interval estimates are not typically highly complex nor impossible to cope with analytically.

Figure 43 graphically demonstrates the fundamentals of the three estimators of a population reliability parameter.

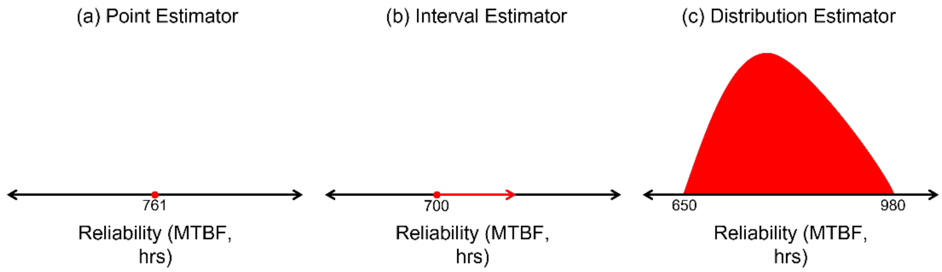


Figure 43. Reliability population parameter estimators.

The point estimate of reliability indicated within Figure 43(a) is of limited value to a decision maker. It infers no information on the true MTBF or the extent of any bias, the sample distribution or the range of variance. As a result, we can only simply state that the reliability expectation is 761 hours MTBF. This is particularly dangerous as a decision maker may internally assess the true reliability as this value when this is unlikely to be true.

Figure 43(b) displays a reliability interval estimator. When compared against the point estimate, the interval estimate provides significantly more information. In this example, a lower confidence limit has been applied. If we were to assume a 90% lower confidence limit has been used, we could state to the decision maker that we are 90% confident that the true reliability is equal to or greater than 700 hours MTBF. However, there exists no indication on the distribution of true MTBF probabilities or if the estimator is distributed symmetrically about the true reliability.

Figure 43(c) demonstrates an example distribution estimator. This permits the conveyance of the complete information regarding our belief or knowledge of the true reliability based on sample test data. For example, we could state that the true reliability lies between 650 and 980 hours MTBF with an expectation of approximately 690 hours. Typically, an interval estimate would also be applied over the distribution estimate to provide additional information to the decision maker.

6.9 *Conclusions*

Communicating reliability growth plan risk to decision makers, both potential risk during plan design and realized risk during plan execution, is a challenging endeavor. Those communicating risks need to understand their audience, determine the appropriate communication strategy, frame their arguments, and ensure clear communication goals. Unfortunately, communicating with decision makers is not typically taught during the formal education of engineers; instead, it is a skill that requires continual development and refinement through experience and practice. This chapter has presented several methods to effectively communicate reliability growth plan risk to decision makers using risk-based focused approaches. Additionally, it has highlighted concerns associated with observations based on point estimates, often used because they are intuitive and easy to understand. The downside of point estimates is that they are frequently erroneous. Even what appears to be a reasonable point estimate is very likely to include error.

Chapter 7: Conclusions, Contributions, and Recommendations

7.1 *Conclusions*

This dissertation presented a new Bayesian approach to undertake reliability growth planning and projection activities for discrete-use systems. A methodology was developed and proposed that supported the use of all available data from previous test and analysis activities. This improved on the approach offered by Hall et al. as the proposed method considers all observed failures rather than only the first observed failure observed on test for any particular mode.

Computational analysis and simulation results supported the development of the proposed method and were used to validate the results and the developed model.

Comparisons were made between the system-level initial reliability estimators using two Bayesian approaches. In many of the cases considered, the Bayesian approach developed in this dissertation outperformed the comparative Bayesian approach offered by Hall et al.

A hypothetical case study based on a surface-to-air missile system was presented that demonstrated the novel way the proposed Bayesian approach may be applied and the utility offered. The case study demonstrated how the approach might be used throughout the entire system development lifecycle, including developmental, acceptance, and operational testing, together with a discussion on the benefits, disadvantages, and constraints that apply.

Finally, methods for effectively communicating reliability growth plans and risks to decision makers were examined and developed within a reliability engineering context.

The research conducted drew the following summarized conclusions.

7.1.1 Utility of a Bayesian Approach

The ability to consider model uncertainty within a single framework, although currently underused, is a significant advantage of Bayesian methods. The research has identified the following concerning the utility of adopting a Bayesian approach to reliability growth planning:

- The main reason for using a Bayesian approach in developing a reliability growth plan is that it facilitates representing and taking fuller account of the uncertainties related to models and parameter values. In contrast, most plans based on maximum likelihood (or least squares) estimation involve fixing the values of parameters that may, and usually do, have an important bearing on the outcome of the analysis and for which there is considerable uncertainty. One of the significant benefits of the proposed Bayesian approach is the ability to incorporate prior information.

- While other reliability growth program development approaches use prior information by specifying levels or ranges of individual parameters for use in sensitivity analysis, the Bayesian approach forces the reliability practitioner to look at historical data sets or to canvass expert knowledge to determine what is known about the system parameters and the test processes. Most traditional growth planning methods do not use any of the quantitative information that could be gathered from historical experience or previous tests and, in effect, treat each system test as a new and independent problem.

- In the past, the effects of uncertainty in developing reliability growth plans have been evaluated through sensitivity analysis. This generally involved changing the value of a single parameter only and rerunning the growth plan assessment. This limitation to a single parameter was due to time constraints and to avoid large amounts of model output. There is a need for sensitivity analysis for any reliability growth program that demonstrates significant uncertainty. However, current practice cannot guarantee that some (reasonably plausible) combination of parameter values does not give rise to behavior that would not be expected from the results of sensitivity tests that involve changing the value of a single parameter only. This is particularly true for 'human unknowns.' Managers will often change their focus partway through a reliability growth plan execution. Decision makers may alter the management strategy applied in correcting observed failure modes, and the effectiveness of corrective actions is varied and never guaranteed with certainty.

7.1.2 Comparison between Hall's and Proposed Model Approaches

Early research identified potential areas that could be improved in developing a new Bayesian model based on Hall's previous work. In researching a new approach, the following was evident when comparing the two methods:

- The approach developed by Hall et al. considers only the first observation of each mode failure. Furthermore, this approach treats both the fix effectiveness factor (FEF) and management strategy (MS) as deterministic mean values. While this approach greatly simplifies the model, it also results in errors in estimating the posterior reliability distribution.

- The proposed approach presents opportunities to apply more significant uncertainty to the various model parameters to ensure that the resultant posterior reliability distribution reasonably reflects the range of possible test outcomes.

- The proposed approach applies the use of the FEF in a slightly different manner than that previously used in discrete-use reliability growth projection models. The FEF has been previously applied by directly scaling the probability of failure for a failure mode. The proposed approach yields a failure mode reliability after corrective action, which provides consistency between continuous-use and discrete-use models.

7.1.3 Simulation in Support of Reliability Growth Program Development

Reliability simulation in support of growth program design and system performance optimization creates a system-level reliability digital twin for consideration. It allows reliability practitioners to simulate any action and optimize their reliability growth plans to achieve the best outcomes within imposed constraints while meeting resource and management requirements. During the research, the following key points were identified:

- Simulation analysis is a powerful tool in modeling the reliability of systems. Proper application requires an understanding of the underlying principles.

- It can be exceptionally intensive computationally to apply Bayes Theorem to complex models. It often takes significant computer time, even on reasonably powerful personal computers, to analyze long-duration test phases with many model variables. The algebraic demands of the methods (including the need for a complete understanding of probability theory) have also discouraged the application of the method more widely. However, to conduct defensible decision analyses for assessments based on maximum likelihood estimation, it is usually necessary to conduct a bootstrap analysis. Although not usually as intensive computationally as applying Bayes Theorem, such an analysis can also often take significant time on a personal computer. Furthermore, even seemingly simple approaches such as bootstrapping are not without their theoretical traps.

7.1.4 Comparison of Bayesian Initial Reliability Estimation Approaches

Although the research did not initially seek to compare Bayesian initial reliability estimation approaches specifically, the following observations were made as a consequence:

- The process of choosing prior distributions can be very time-consuming and frustrating. Reliability practitioners may provide prior distributions which are either inconsistent or far too precise. Although expert opinion is often the dominant method for determining priors and is the source of many problems, it is expected that prior distributions will increasingly be determined by analysis of information from synthesis studies and hence depend less on unreliable expert opinion.

- The majority of the problems encountered during the development of Bayesian assessments have resulted from arguments about the choice of prior distributions. The studies comparing the utility of both Hall's and the proposed prior methodologies have demonstrated that these approaches are valid and can be conducted with minimal data. This research has shown that both methods perform well, with minimal cumulative relative error observed compared to the system's true initial reliability.

- The proposed method outperforms Hall's approach in many cases.

7.1.5 Effectively Communicating Reliability Growth Plans and Risk

The research did not specifically seek to consider the methods and means that reliability practitioners communicate reliability growth plans and risk. The author's frustrations in reliability engineering communication with senior decision makers resulted in its inclusion, including presenting a conference paper and the invitation to submit an article for publishing consideration to a peer-reviewed journal. The following summarized engineering communication research observations are presented:

- The global communication landscape is rapidly changing, and effectively communicating reliability engineering information can be derailed by many factors.
- It is becoming more frequent that engineers face misinformation, purposeful denial, ignorance or indifference, political expedience, and many other problems.
- The reliability engineering environment has become increasingly complex and reflective of current society. Within the community, diminishing trust in institutions and experts, increasing use of social media as an information source, and the rise of alternative viewpoints inconsistent with evidence are commonplace.
- Most reliability practitioners are adept in a range of "hard" or quantitative skills. Conversely, the stereotypical engineer is considered deficient in "soft" or professional skills.
- Effectively communicating reliability engineering messages to decision makers necessitates that the reliability practitioner:
    - know the audience,

- understand the audience and their likely reaction to new information by conducting audience segmentation analyses,
- focusing on a key messaging strategy reflective of the audience needs,
- simplify the message without removing too much detail,
- engage with the audience through storytelling to build rapport,
- understand the goals of the communication, and
- frame the communication appropriately so that it appeals to the audience.

7.2 *Contributions*

The major contributions of this work are as follows:

1. A new Bayesian reliability projection model was proposed that considers the uncertainty surrounding discrete-use systems under arbitrary corrective action regimes to address failure modes. This differs from current models that fail to address the arbitrary nature of corrective action application strategies observed in real-world test situations. Additionally, the proposed strategy permits a probabilistic assessment of the test program, accounting for uncertainty in initial reliability and management variables.
2. An extension to the proposed Bayesian discrete-use projection model was developed by considering developmental, acceptance, and operational testing

aspects through simulation of failures and corrective actions. This allows the formulation of a holistic reliability growth plan framework that encompasses the entire system lifecycle. The approach considers the posterior distribution from each phase of developmental testing as the prior for the following growth test event. The same methodology is employed using the posterior from the final phase of reliability growth testing as the acceptance testing prior. It then follows that the acceptance testing posterior distribution forms the prior for subsequent operational testing. Importantly the approach is flexible enough to permit the combination of data from any test activity conducted in any order. The approach reduces unrealistic and unattainable reliability testing that may result from a purely statistical analysis. The proposed methodology also permits planning for combined developmental and operational test activities within a financially constrained context.

3. The research presented an approach for combining disparate data from various sources to establish prior distributions on system reliability.

4. The research developed and presented novel methods to assist reliability engineers in communicating developmental system reliability growth plans and risks to decision makers more effectively. The research takes essential facets of communication theory from marketing, management, business, and advertising and adapts them to the reliability engineering endeavor.

7.3 *Recommendations for Future Research*

Based on the results of this research and the knowledge gaps and opportunities consequently identified, recommendations for future research include:

- The methodology advanced in this research is aimed to describe reliability growth of discrete-use. Estimating the reliability of these systems presents a significant challenge due to their serial structure. These systems invariably operate sequentially by their design nature, meaning that the current state of operation depends solely on the preceding stages. This issue of failure mode "masking" has been recognized, modeled, and described elsewhere. To minimize the preclusion influence, one possibility could be to apply the planning model approach separately to each stage of system function. It may be necessary to assign stage developmental reliability goals that need to be achieved at the end of the reliability growth program. This assumes that the system comprising statistically independent stages and given proper inputs from the preceding stages can be serially decomposed.

- The assumption that failure probabilities for each mode are drawn from the same fixed population distribution may require further assessment and understanding. In real-world systems, a significant difference in failure mode probabilities within a system is likely to exist. Some modes appear more often due to design inadequacies that result in a higher probability of occurrence. These can often be quickly identified and corrective actions implemented. Other modes may be observed less frequently and have a lower probability of occurrence. Consequently, the assumption that mode failure probabilities may be drawn independently from the same fixed distribution should be challenged further.

- The development of management metrics similar to those presented by Hall et al. [33] could contribute significantly to the available literature on the subject. There are several categories of metrics that could be considered for future utility:
    - those that focus on the different failure modes,
    - those that focus on reliability, and
    - those that focus on the costs and benefits of either choosing to address failure modes via the application of corrective actions or not.
- The development of management metrics that promote an understanding of failure modes, both observed and hidden, may help guide resource allocation for future investigations into newly discovered modes.
- The development of Bayesian derived reliability-focused metrics are likely to be central in enlightening decision makers in achieving better outcomes.

- The development of cost-based metrics could assist in guiding decision makers by highlighting how resources are spent on the improved assessment and reliability of the system population.

- The intoduction and assessment of Bayesian heirarchical models for families of discrete-use systems whose designs are similar in nature under the assumption of interchangeability. Bayesian hierarchical models may be suitable to allow partial information to be leveraged among different products. Since systems from the same family have very close reliability estimation, it is reasonable to build a Bayesian hierarchical model that can partially leverage the reliabilities of previous products as prior information. Thus, when estimating the reliability of a newly released product or system, partial strength can be borrowed from similar products to reduce the uncertainty of estimated reliability. An example could be an armored fighting vehicle turret and its enclosed primary and secondary weapon systems. The same turret and weapons may be utilized on different tracked and wheeled vehicle systems.

- The possible extension of the Bayesian modeling and simulation concept application to complex mixed continuous-use and discrete-use systems. The modeling of complex systems, which are incredibly complex and contain mixed continuous-use and discrete-use subsystems, could significantly benefit through the adoption of

Bayesian methodologies for system-level reliability growth planning. Examples of highly complex mixed continuous-use and discrete-use systems include surface and subsurface warships, aeronautical and space vehicles, armored fighting vehicles, and communications systems.

7.4 *Resulting Publications*

This research has resulted in several conference presentations, papers, and journal articles. The following sections list the published work as a direct result of the research conducted. Where available at dissertation finalization and submission, links to the described documents have been included.

7.4.1 Conference Presentations and Papers

The research findings have been presented at the following conferences through the submission of conference papers:

P. J. Nation and M. Modarres, "*Modelling uncertainty in reliability growth planning for continuous-use systems*," presented at Systems Engineering Test and Evaluation Conference, Canberra, Australia, April 29 to May 1, 2019, https://beta.informit.org/doi/abs/10.3316/informit.443936245043201.

P. J. Nation and M. Modarres, "*Modelling continuous-use system reliability growth utilising disparate source data*," presented at Systems Engineering Test and Evaluation Conference, Canberra, Australia, April 29 to May 1, 2019, https://beta.informit.org/doi/10.3316/informit.443954878014460.

P. J. Nation, M. Wayne and M. Modarres, "*A Bayesian approach to complex discrete-use system reliability growth planning under delayed or arbitrary corrective actions*," presented at Integrated Project Engineering Congress, Online, May 26-28, 2021.

P. J. Nation, M. Wayne and M. Modarres, "*A comparison of empirical Bayes hyperparameter approaches for discrete-use system initial reliability estimation*," presented at Integrated Project Engineering Congress, Online, May 26-28, 2021.

P. J. Nation, M. Wayne and M. Modarres, "*Effectively communicating developmental system reliability growth plans and risk to decision makers*," presented at Integrated Project Engineering Congress, Online, May 26-28, 2021.

P. J. Nation, M Wayne and M Modarres, "*A Bayesian approach to complex discrete-use system reliability growth planning under delayed or arbitrary corrective actions*," submitted at International Mechanical Engineering Congress and Exposition, Online, November 1-5, 2021 (awarded honorable mention in the American Society of Mechanical Engineers Safety Engineering, Risk, and Reliability Analysis Division's Student Safety Innovation Challenge).

7.4.2 Journal Articles

The research conducted has resulted in the submission of the following peer-reviewed professional journal articles:

P. J. Nation and M. Modarres, "*Modelling uncertainty in reliability growth planning for continuous-use systems utilising disparate source data*," Australian Journal of Multi-Disciplinary Engineering, Vol 15, Issue 1, 2019, pp.2-16, DOI: 10.1080/14488388.2019.1661808.

P. J. Nation, M. Wayne and M. Modarres, "*Effectively communicating developmental system reliability growth plans and risk*," submitted for consideration upon invitation to the Australian Journal of Multi-Disciplinary Engineering.

P. J. Nation, M. Wayne and M. Modarres, "*A Bayesian approach to discrete-use system reliability growth planning under delayed or arbitrary corrective actions*," submitted for consideration to the Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability.

Appendix 1: Mathematical Derivations

This appendix contains the mathematical derivations for the empirical Bayes initial reliability estimator approach detailed in Chapter 3 and demonstrated in Chapter 4.

Note that many of the equations have been included on individual pages rather than breaking them over multiple pages to improve readability.

Empirical Bayes Estimates for Discrete Projection

The marginal likelihood for a single failure mode with $n_i = (n_{i,1}, n_{i,2})$ is given as

$$L(n_i) = \sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{\Gamma(a+n_{i,1})\Gamma\left(b+v_i-n_{i,1}+(1-d_i)(T-v_i-n_{i,2}+j)\right)}{\Gamma\left(a+b+v_i+(1-d_i)(T-v_i-n_{i,2}+j)\right)}$$

$$= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j \frac{\Gamma(a+n_{i,1})\Gamma(b+v_i-n_{i,1}+\tau_{i,j})}{\Gamma(a+b+v_i+\tau_{i,j})} \quad (1)$$

where $\tau_{i,j} = (1-\mathrm{d_i})(\mathrm{T}-\mathrm{v_i}-\mathrm{n_{i,2}}+\mathrm{j})$.

The total likelihood over all modes is then

$$\mathrm{L}(n) = \sum_{i=1}^{k} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j \frac{\Gamma(a+n_{i,1})\Gamma(b+v_i-n_{i,1}+\tau_{i,j})}{\Gamma(a+b+v_i+\tau_{i,j})} \quad (2)$$

The log-likelihood is

$$\begin{aligned}
\ell(\mathrm{n}) &= \log \mathrm{L}(n) \\
&= \sum_{i=1}^{k} \Bigg[\log \Gamma(a+b) - \log \Gamma(a) - \log \Gamma(b) \\
&\qquad + \log \left(\sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j \frac{\Gamma(a+n_{i,1}) \Gamma(b+v_i-n_{i,1}+\tau_{i,j})}{\Gamma(a+b+v_i+\tau_{i,j})} \right) \Bigg] \\
&= \sum_{i=1}^{k} \Bigg[\log \Gamma(a+b) - \log \Gamma(a) - \log \Gamma(b) + \log \Gamma(a+n_{i,1}) \\
&\qquad + \log \left(\sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j \frac{\Gamma(b+v_i-n_{i,1}+\tau_{i,j})}{\Gamma(a+b+v_i+\tau_{i,j})} \right) \Bigg]
\end{aligned} \tag{3}$$

Taking the partial derivative WRT a yields

$$\frac{\partial \ell(\mathrm{n})}{\partial a}$$

$$= \sum_{i=1}^{k} \left[\psi(a+b) - \psi(a) + \psi\left(a + n_{i,1}\right) \right.$$

$$+ \left(\frac{1}{\sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j \frac{\Gamma\left(b + v_i - n_{i,1} + \tau_{i,j}\right)}{\Gamma\left(a + b + v_i + \tau_{i,j}\right)}} \right.$$

$$\left. \left. \cdot \sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^{j+1} \frac{\Gamma\left(b + v_i - n_{i,1} + \tau_{i,j}\right)}{\Gamma\left(a + b + v_i + \tau_{i,j}\right)^2} \Gamma'\left(a + b + v_i + \tau_{i,j}\right) \right) \right]$$

$$= \sum_{i=1}^{k} \left[\psi\left(a + n_{i,1}\right) - \psi(a) \right] + \sum_{i=1}^{k} \left[\psi(a+b) + \frac{\sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^{j+1} \frac{\Gamma\left(b + v_i - n_{i,1} + \tau_{i,j}\right)}{\Gamma\left(a + b + v_i + \tau_{i,j}\right)}}{\sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^j \frac{\Gamma\left(b + v_i - n_{i,1} + \tau_{i,j}\right)}{\Gamma\left(a + b + v_i + \tau_{i,j}\right)}} \right]$$

$$\cdot \psi\left(a + b + v_i + \tau_{i,j}\right) \tag{4}$$

Similarly, taking the partial derivative WRT b gives

$$\frac{\partial \ell(\mathrm{n})}{\partial b}$$

$$= \left[\sum_{i=1}^{k} \psi(a+b) - \psi(b) + \frac{1}{\sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^{j} \frac{\Gamma(b+v_i-n_{i,1}+\tau_{i,j})}{\Gamma(a+b+v_i+\tau_{i,j})}} \right.$$

$$\cdot \sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^{j+1}$$

$$\left. \cdot \left(\frac{\Gamma'(b+v_i-n_{i,1}+\tau_{i,j})\Gamma(a+b+v_i+\tau_{i,j}) - \Gamma(b+v_i-n_{i,1}+\tau_{i,j})\Gamma'(a+b+v_i+\tau_{i,j})}{\Gamma(a+b+v_i+\tau_{i,j})^2} \right) \right] \quad (5)$$

Breaking the likelihood into observed and unobserved parts (m! ways to arrange observed modes with $\binom{k}{m}$ choices) yields

$$\mathrm{L}(n) = \mathrm{m}! \binom{k}{m} \left[\frac{\Gamma(a+b)\Gamma(a+n_{i,1})}{\Gamma(a)\Gamma(b)} \sum_{j=0}^{n_{i,2}} \binom{n_{i,2}}{j} (-1)^{j+1} \frac{\Gamma(b+v_i-n_{i,1}+\tau_{i,j})}{\Gamma(a+b+v_i+\tau_{i,j})} \right]$$

$$\cdot \left[\frac{\Gamma(a+b)\cancel{\Gamma(a)}\Gamma(b+T)}{\cancel{\Gamma(a)}\Gamma(b)\Gamma(a+b+T)} \right]^{k-m} \quad (6)$$

Given that $\tilde{\mathrm{n}} = \mathrm{a} + \mathrm{b}$ and $\mathrm{a} = \tilde{\mathrm{n}}(1 - \mathrm{R_I})^{1/\mathrm{k}}$ then

$$\mathrm{b} = \tilde{\mathrm{n}} - \mathrm{a} = \tilde{\mathrm{n}} - \tilde{\mathrm{n}}(1 - R_I)^{1/k} = \tilde{n}\left(\cancel{1} - \cancel{1} + R_I{}^{1/k}\right) = \tilde{n} R_I^{1/k} \quad (7)$$

For $\ell(\mathrm{n}) = \log \mathrm{L}(n)$, the unobserved part then becomes

$$\ell_{unobs}(\mathrm{n}) = (k-m)\log\left[\frac{\Gamma(\tilde{n})\Gamma\left(\tilde{n}R_I^{1/k} + T\right)}{\Gamma\left(\tilde{n}R_I^{1/k}\right)\Gamma(\tilde{n}+T)}\right]$$

$$= \left(1-\frac{m}{k}\right)k\left[\log\Gamma(\tilde{n}) + \log\Gamma\left(\tilde{n}R_I^{1/k} + T\right) - \log\Gamma\left(\tilde{n}R_I^{1/k}\right) - \log\Gamma(\tilde{n}+T)\right] \qquad (8)$$

Taking the limit as $k \to \infty$

$$\lim_{k\to\infty}\left(1-\frac{m}{k}\right)k\left[\log\Gamma(\tilde{n})+\log\Gamma\left(\tilde{n}R_I^{1/k}+T\right)-\log\Gamma\left(\tilde{n}R_I^{1/k}\right)-\log\Gamma(\tilde{n}+T)\right]$$

$$=\lim_{k\to\infty}\frac{\log\Gamma(\tilde{n})-\log\Gamma(\tilde{n}+T)+\log\dfrac{\Gamma\left(\tilde{n}R_I^{1/k}+T\right)}{\Gamma\left(\tilde{n}R_I^{1/k}\right)}}{1/k}$$

$$=\lim_{k\to\infty}\frac{\log\Gamma(\tilde{n})-\log\Gamma(\tilde{n}+T)+\log\dfrac{\left(\tilde{n}R_I^{1/k}+T-1\right)!}{\left(\tilde{n}R_I^{1/k}-1\right)!}}{1/k}$$

$$=\lim_{k\to\infty}\frac{\log\Gamma(\tilde{n})-\log\Gamma(\tilde{n}+T)+\log\left[\left(\tilde{n}R_I^{1/k}+T-1\right)\left(\tilde{n}R_I^{1/k}+T-2\right)\cdots\tilde{n}R_I^{1/k}\right]}{1/k}$$

$$=\lim_{k\to\infty}\frac{\log\Gamma(\tilde{n})-\log\Gamma(\tilde{n}+T)+\sum_{j=0}^{T-1}\log\left(\tilde{n}R_I^{1/k}+j\right)}{1/k}$$

$$=\lim_{k\to\infty}\frac{\sum_{j=0}^{T-1}\dfrac{1}{\tilde{n}R_I^{1/k}+j}-\tilde{n}\log R_I\,R_I^{1/k}\left(\cancel{-1/k^2}\right)}{\cancel{-1/k^2}}$$

$$=\lim_{k\to\infty}\tilde{n}R_I^{1/k}\log R_I\sum_{j=0}^{T-1}\frac{1}{\tilde{n}R_I^{1/k}+j}$$

$$=\tilde{n}\log R_I\sum_{j=0}^{T-1}\frac{1}{\tilde{n}+j}$$

$$=\log R_I^{\tilde{n}}\left[\psi(\tilde{n}+T)-\psi(\tilde{n})\right] \tag{9}$$

Similarly, for the observed likelihood

$$\ell_{obs}(n)$$

$$= \log\left[\frac{k!}{(k-m)!}\prod_{i=1}^{m}\frac{\Gamma(\tilde{n})\Gamma\left(\tilde{n}\left(1-R_I^{1/k}\right)+n_{i,1}\right)}{\Gamma\left(\tilde{n}\left(1-R_I^{1/k}\right)\right)\Gamma\left(\tilde{n}R_I^{1/k}\right)}\sum_{j=0}^{n_{i,2}}\frac{\Gamma\left(\tilde{n}R_I^{1/k}+v_i-n_{i,1}+\tau_{i,j}\right)}{\Gamma\left(\tilde{n}+v_i+\tau_{i,j}\right)}\binom{n_{i,2}}{j}(-1)^j\right]$$

$$= \log\left[\prod_{i=1}^{m}\frac{(k+1-1)\Gamma(\tilde{n})\Gamma\left(\tilde{n}\left(1-R_I^{1/k}\right)+n_{i,1}\right)}{\Gamma\left(\tilde{n}\left(1-R_I^{1/k}\right)\right)\Gamma\left(\tilde{n}R_I^{1/k}\right)}\sum_{j=0}^{n_{i,2}}\frac{\Gamma\left(\tilde{n}R_I^{1/k}+v_i-n_{i,1}+\tau_{i,j}\right)}{\Gamma\left(\tilde{n}+v_i+\tau_{i,j}\right)}\binom{n_{i,2}}{j}(-1)^j\right]$$

$$= \sum_{i=1}^{m}\log\left[\frac{\left(1+{}^{1}/_{k}-{}^{i}/_{k}\right)\left(\tilde{n}\left(1-R_I^{1/k}\right)+n_{i,1}-1\right)\left(\tilde{n}\left(1-R_I^{1/k}\right)+n_{i,1}-2\right)\cdots\left(\tilde{n}\left(1-R_I^{1/k}\right)+n_{i,1}-1\right)k\tilde{n}\left(1-R_I^{1/k}\right)\Gamma(\tilde{n})}{\Gamma\left(\tilde{n}R_I^{1/k}\right)}\right.$$

$$\left.\cdot\sum_{j=0}^{n_{i,2}}\frac{\Gamma\left(\tilde{n}R_I^{1/k}+v_i-n_{i,1}+\tau_{i,j}\right)}{\Gamma\left(\tilde{n}+v_i+\tau_{i,j}\right)}\binom{n_{i,2}}{j}(-1)^j\right] \tag{10}$$

From the projection model development $\lim_{k\to\infty} k\tilde{n}\left(1-R_I^{1/k}\right) = -\tilde{n}\log R_I$, therefore

$$\lim_{k\to\infty}\ell_{obs}(n) = \sum_{i=1}^{m}\log\left[\frac{\cancel{\Gamma(\tilde{n})}\left(n_{i,1}-1\right)\left(n_{i,1}-2\right)\left(n_{i,1}-3\right)\cdots(1)\tilde{n}(-\tilde{n}\log R_I)}{\cancel{\Gamma(\tilde{n})}}\right.$$

$$\left.\cdot\sum_{j=0}^{n_{i,2}}\frac{\Gamma\left(\tilde{n}+v_i-n_{i,1}+\tau_{i,j}\right)}{\Gamma\left(\tilde{n}+v_i+\tau_{i,j}\right)}\binom{n_{i,2}}{j}(-1)^j\right]$$

$$= \sum_{i=1}^{m}\left[\log\prod_{j=1}^{n_{i,1}-1} j + \log(-\tilde{n}\log R_I) + \log\sum_{j=0}^{n_{i,2}}\frac{\Gamma\left(\tilde{n}+v_i-n_{i,1}+\tau_{i,j}\right)}{\Gamma\left(\tilde{n}+v_i+\tau_{i,j}\right)}\binom{n_{i,2}}{j}(-1)^j\right] \tag{11}$$

The limiting form of the likelihood is then

$$
\begin{aligned}
\ell_\infty(n) &= \lim_{k\to\infty} \ell(n) \\
&= \sum_{i=1}^{m}\left[\sum_{j=1}^{n_{i,1}-1} \log j + \log\left(\sum_{j=0}^{n_{i,2}} \frac{\Gamma\left(\tilde{n}+v_i-n_{i,1}+\tau_{i,j}\right)}{\Gamma\left(\tilde{n}+v_i+\tau_{i,j}\right)}\binom{n_{i,2}}{j}(-1)^j\right)\right] \\
&\qquad + m\log(-\tilde{n}\log R_I) + \tilde{n}\log R_I\,[\psi(\tilde{n}+T)-\psi(\tilde{n})] \\
&= \sum_{i=1}^{m}\left[\sum_{j=1}^{n_{i,1-1}} \log j + \log\left(\sum_{j=0}^{n_{i,2}} \frac{1}{\left(\tilde{n}+v_i+\tau_{i,j}-1\right)\cdots\left(\tilde{n}+v_i+\tau_{i,j}\right)}\binom{n_{i,2}}{j}(-1)^j\right)\right] \\
&\qquad + m\log(-\tilde{n}\log R_I) + \tilde{n}\log R_I\,[\psi(\tilde{n}+T)-\psi(\tilde{n})] \\
&= \sum_{i=1}^{m}\left[\sum_{j=1}^{n_{i,1-1}} \log j + \log\left(\sum_{j=0}^{n_{i,2}} \prod_{q=1}^{n_{i,2}} \frac{1}{\tilde{n}+v_i+\tau_{i,j}-q}\binom{n_{i,2}}{j}(-1)^j\right)\right] + \mathrm{m}\log(-\tilde{n}\log R_I) \\
&\qquad + \tilde{n}\log R_I\,[\psi(\tilde{n}+T)-\psi(\tilde{n})]
\end{aligned}
\tag{12}
$$

IOT find the partial derviative of $\ell_\infty(n)$ WRT $\tilde{n}$, let

$$
\begin{aligned}
\mathrm{f} &= \left[\left(\tilde{n}+v_i+\tau_{i,j}-1\right)\left(\tilde{n}+v_i+\tau_{i,j}-2\right)\cdots\left(\tilde{n}+v_i+\tau_{i,j}-n_{i,1}\right)\right]^{-1} \\
&= \frac{\cancel{\tilde{n}+v_i+\tau_{i,j}}}{\left(\tilde{n}+v_i+\tau_{i,j}-n_{i,1}\right)\left(\tilde{n}+v_i+\tau_{i,j}-n_{i,1}+1\right)\cdots\left(\tilde{n}+v_i+\tau_{i,j}-\cancel{n_{i,1}}+\cancel{n_{i,1}}-1\right)\left(\cancel{\tilde{n}+v_i+\tau_{i,j}}-\cancel{n_{i,1}}+\cancel{n_{i,1}}\right)}
\end{aligned}
\tag{13}
$$

The partial derivative of f WRT $\tilde{n}$, noting the first line of Equation (14) is derived from Mathematica output) is

$$\frac{\partial f}{\partial \tilde{n}} = \frac{\tilde{n} + v_i + \tau_{i,j}}{\left(\tilde{n} + v_i + \tau_{i,j}\right)^2 \prod_{q=0}^{n_{i,1}}\left(\tilde{n} + v_i + \tau_{i,j} - n_{i,1} + 1 + q\right)}$$

$$+ \frac{1 - \left(\tilde{n} + v_i + \tau_{i,j}\right)\psi\left(\tilde{n} + v_i + \tau_{i,j} + 1\right) - \psi\left(\tilde{n} + v_i + \tau_{i,j} - n_{i,1} + 1\right)}{\left(\tilde{n} + v_i + \tau_{i,j} - n_{i,1}\right)\prod_{q=0}^{n_{i,1}}\left(\tilde{n} + v_i + \tau_{i,j} - n_{i,1} + 1 + q\right)}$$

$$= -\left[\left(\tilde{n} + v_i + \tau_{i,j} - 1\right)\left(\tilde{n} + v_i + \tau_{i,j} - 2\right)\cdots\left(\tilde{n} + v_i + \tau_{i,j} - n_{i,1}\right)\right]^{-2}$$

$$\cdot\left[\prod_{q=2}^{n_{i,1}} \tilde{n} + v_i + \tau_{i,j} - q + \prod_{\substack{q=1 \\ q\neq 2}}^{n_{i,1}} \tilde{n} + v_i + \tau_{i,j} - q + \cdots \right.$$

$$\left. + \prod_{q=1}^{n_{i,1}-1} \tilde{n} + v_i + \tau_{i,j} - q\right]$$

$$= -\left[\frac{\prod_{q=2}^{n_{i,1}} \tilde{n} + v_i + \tau_{i,j} - q}{\left(\prod_{q=1}^{n_{i,1}} \tilde{n} + v_i + \tau_{i,j} - q\right)^2} + \frac{\prod_{\substack{q=1 \\ q\neq 2}}^{n_{i,1}} \tilde{n} + v_i + \tau_{i,j} - q}{\left(\prod_{q=1}^{n_{i,1}} \tilde{n} + v_i + \tau_{i,j} - q\right)^2} + \cdots \right.$$

$$\left. + \frac{\prod_{q=1}^{n_{i,1}-1} \tilde{n} + v_i + \tau_{i,j} - q}{\left(\prod_{q=1}^{n_{i,1}} \tilde{n} + v_i + \tau_{i,j} - q\right)^2}\right]$$

$$= -\left[\frac{1}{\left(\tilde{n} + v_i + \tau_{i,j} - 1\right)\prod_{q=1}^{n_{i,1}} \tilde{n} + v_i + \tau_{i,j} - q} \right.$$

$$+ \frac{1}{\left(\tilde{n} + v_i + \tau_{i,j} - 2\right)\prod_{q=1}^{n_{i,1}} \tilde{n} + v_i + \tau_{i,j} - q} + \cdots$$

$$\left. + \frac{1}{\left(\tilde{n} + v_i + \tau_{i,j} - n_{i,1}\right)\prod_{q=1}^{n_{i,1}} \tilde{n} + v_i + \tau_{i,j} - q}\right]$$

$$= -\left(\frac{1}{\prod_{q=1}^{n_{i,1}} \tilde{n} + v_i + \tau_{i,j} - q}\right)\left[\frac{1}{\tilde{n} + v_i + \tau_{i,j} - 1} + \frac{1}{\tilde{n} + v_i + \tau_{i,j} - 2} + \cdots \right.$$

$$\left. + \frac{1}{\tilde{n} + v_i + \tau_{i,j} - n_{i,1}}\right] \qquad (14)$$

(continued next page)

$$\frac{\partial f}{\partial \tilde{n}} = -\left(\frac{1}{\prod_{q=1}^{n_{i,1}} \tilde{n} + v_i + \tau_{i,j} - q}\right)\sum_{q=1}^{n_{i,1}} \frac{1}{\tilde{n} + v_i + \tau_{i,j} - q} \tag{14}$$

The partial dervivative of $\ell_\infty(n)$ WRT $\tilde{n}$ is then

$$\begin{aligned}
&\frac{\partial \ell_\infty(n)}{\partial \tilde{n}} \\
&= \sum_{i=1}^{m}\left\{\left(\frac{1}{\sum_{j=0}^{n_{i,2}} \prod_{q=1}^{n_{i,1}} \frac{1}{\tilde{n} + v_i + \tau_{i,j} - q}}\right)\sum_{j=0}^{n_{i,2}}\left[\left(\prod_{q=1}^{n_{i,1}} \frac{-1}{\tilde{n} + v_i + \tau_{i,j} - q}\right)\sum_{q=1}^{n_{i,1}} \frac{1}{\tilde{n} + v_i + \tau_{i,j} - q}\right]\right\} \\
&+ \frac{\cancel{-}m}{\tilde{n}\,\cancel{\log R_I}}\left(\cancel{-\log R_I}\right) + \log R_I\left[\psi(\tilde{n}+T) - \psi(\tilde{n})\right] + \tilde{n}\log R_I\left[\psi'(\tilde{n}+T) - \psi'(\tilde{n})\right]
\end{aligned} \tag{15}$$

The partial dervivative of $\ell_\infty(n)$ WRT R_I is

$$\frac{\partial \ell_\infty(n)}{\partial R_I} = \frac{\cancel{-}m}{\cancel{\tilde{n}}\log R_I}\left(\frac{\cancel{-\tilde{n}}}{R_I}\right) + \frac{\tilde{n}}{R_I}\left[\left[\psi(\tilde{n}+T) - \psi(\tilde{n})\right]\right] \tag{16}$$

Let the partial derivative $\ell_\infty(n)$ WRT R_I equal zero, then

$$\begin{aligned}
\frac{\partial \ell_\infty(n)}{\partial R_I} = 0 \quad &\Rightarrow \quad \frac{-m}{\cancel{R_I}\log R_I} = \frac{\tilde{n}}{\cancel{R_I}}\left[\psi(\tilde{n}+T) - \psi(\tilde{n})\right] \\
&\Rightarrow \quad \boxed{\log R_I = exp\left[\frac{-m}{\tilde{n}\left[\psi(\tilde{n}+T) - \psi(\tilde{n})\right]}\right]}
\end{aligned} \tag{17}$$

Subsituting the expression for R_I from Equation (17) into $\frac{\partial \ell_\infty(n)}{\partial \tilde{n}}$ from Equation (15) yeields

$$-\sum_{i=1}^{m} \frac{\sum_{j=0}^{n_{i,2}} \left(\prod_{q=1}^{n_{i,1}} \frac{1}{\tilde{n} + v_i + \tau_{i,j} - q}\right) \sum_{q=1}^{n_{i,1}} \frac{1}{\tilde{n} + v_i + \tau_{i,j} - q}}{\sum_{j=0}^{n_{i,2}} \prod_{q=1}^{n_{i,1}} \frac{1}{\tilde{n} + v_i + \tau_{i,j} - q}} + \frac{\cancel{m}}{\cancel{\tilde{n}}}$$

$$- \frac{\cancel{m}}{\cancel{\tilde{n}}[\cancel{\psi(\tilde{n}+T)} - \cancel{\psi(\tilde{n})}]}[\cancel{\psi(\tilde{n}+T)} - \cancel{\psi(\tilde{n})}]$$

$$+ \cancel{\tilde{n}} \frac{-m}{\cancel{\tilde{n}}\big[[\psi(\tilde{n}+T) - \psi(\tilde{n})]\big]}[\psi'(\tilde{n}+T) - \psi'(\tilde{n})]$$

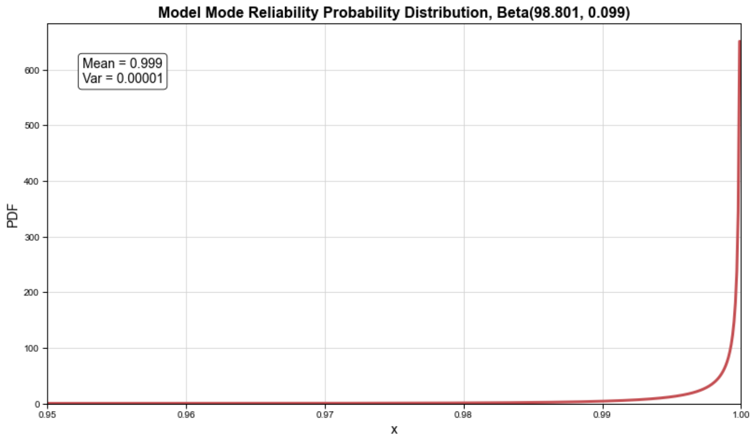
$$\Rightarrow \boxed{\sum_{i=1}^{m} \frac{\sum_{j=0}^{n_{i,2}} \left(\prod_{q=1}^{n_{i,1}} \frac{1}{\tilde{n} + v_i + \tau_{i,j} - q}\right) \sum_{q=1}^{n_{i,1}} \frac{1}{\tilde{n} + v_i + \tau_{i,j} - q}}{\sum_{j=0}^{n_{i,2}} \prod_{q=1}^{n_{i,1}} \frac{1}{\tilde{n} + v_i + \tau_{i,j} - q}} = m \frac{[\psi'(\tilde{n}) - \psi'(\tilde{n}+T)]}{[\psi(\tilde{n}+T) - \psi(\tilde{n})]}} \qquad (18)$$

Appendix 2: Python Programing Language Code

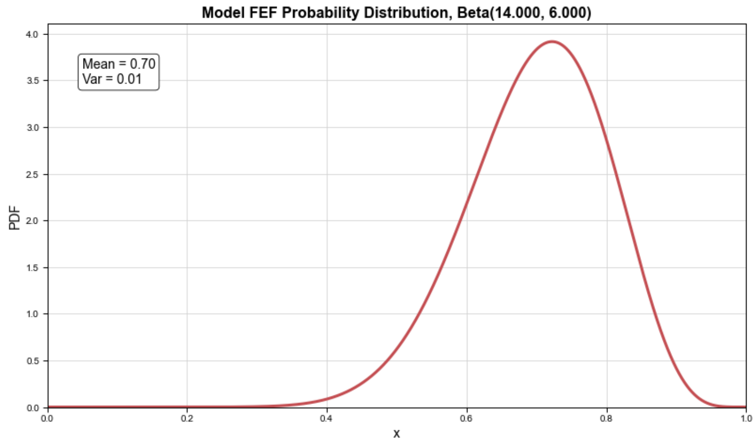
The following Python code was utilized in assessing many of the methodologies proposed in this dissertation. It is included for completeness and to ensure others may replicate and expand upon this work in the future.

```
#########################
# Import libraries
import numpy as np
import math as ma
from mpmath import *
mp.dps = 50
import scipy.stats as ss
import random as rand
from scipy.special import comb, polygamma, psi, btdtri
import scipy.optimize as so
from datetime import datetime
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline

# Set model and simulation parameters
k = 250 # Number of system failure modes
T = 600 # Test demands
pm = 0.001 # Mean mode failure probability
pv = 0.00001 # Mode failure probability variance
pfix = 0.1 # Probability of mode fix during test
c = 0.7 # FEF mean
e = 0.01 # FEF variance
CL = 0.8 # Confidence limit
niter = 1000 # Number of iterations
case = 1 # Case number under examination

# Initialize plotting parameters
sns.set_theme(context='paper', # Set seaborn params
        style='ticks',
        font='sans-serif',
        font_scale=1.4,
        rc={'figure.facecolor': 'white',
          'figure.titlesize': 16,
          'figure.titleweight': 'bold',
          'axes.edgecolor': 'black',
          'axes.facecolor': 'white',
          'axes.labelcolor': 'black',
          'axes.grid': True,
          'grid.color': 'lightgrey',
          'xtick.direction': 'out',
          'ytick.direction': 'out',
          'xtick.color': 'black',
          'ytick.color': 'black',
          'text.color': 'black',
          'axes.titlesize': 16,
          'axes.titleweight': 'bold',
          'axes.labelsize': 14,
          'font.size': 14,
          'xtick.labelsize': 10,
          'ytick.labelsize':10,
          'lines.color': 'black',
          'legend.fontsize':11,
          'legend.loc': 'upper center',
```

```
                'legend.edgecolor': 'white',
                'legend.facecolor': 'white'})

bbox_args = dict(boxstyle='round', fc='white', ec='black', zorder=10)

width, height = 13.333, 7.5 # Set plot width and height params

save_plots = 'Y' # Save plots Y or N

rand_int = 1 # Select number of random simulation iterations to plot on the expected number of failure modes curve

rand_sims = [] # Empty list of random simulation mode failures FOT

randomlist = [] # Select random simulations
for i in range(rand_int):
   n = rand.randint(0, niter)
   randomlist.append(n)

print(randomlist) # List of random simulations for closer examination

>>> [325]

# Calculate Beta(n,x) parameters for mode failure probability distribution and plot
Beta_x = ((1 - pm) / pv - (1 / pm)) * pm**2 # True Beta distribution alpha parameter
Beta_n = Beta_x * ((1 / pm) - 1) # True Beta distribution Beta parameter
n_approx = Beta_x + Beta_n # n tilde for the Beta distribution (used to constrain simulation)
print('Beta(n,x) n parameter for mode failure probability distribution: ' + str('{0:.3f}'.format(Beta_n)))
print('Beta(n,x) x parameter for mode failure probability distribution: ' + str('{0:.3f}'.format(Beta_x)))
print('Beta(n,x) n tilde parameter for mode failure probability distribution: ' + str('{0:.0f}'.format(n_approx)))

fig, ax = plt.subplots(figsize=(width, height)) # Set plot size

# Plot initial mode reliability probability distributions
x = np.linspace(0, 1, 10000)
plt.plot(x, ss.Beta.pdf(x, Beta_n, Beta_x), 'r-', lw=3)

# Control x and y limits
ax.set(ylim=(0.0, None))
ax.set(xlim=(0.95, 1.0))

# Annotate mode failure probability parameter
plt.annotate('Mean = ' + str('{0:.3f}'.format(1 - pm)) + '\n'
          + 'Var = ' + str('{0:.5f}'.format(pv)), xy=(0.05, 0.91), xycoords='axes fraction',
          horizontalalignment='left', verticalalignment='top', bbox=bbox_args)

# Set plot label and axis labels
ax.set(title='Model Mode Reliability Probability Distribution'
      + ', Beta(' + str('{0:.3f}'.format(Beta_n)) + ', ' + str('{0:.3f}'.format(Beta_x)) +')')
ax.set(xlabel='x', ylabel='PDF')

# Save plot as .jpg ready for MS PowerPoint widescreen presentation
if save_plots == 'Y':
   plt.savefig('01 Model Mode reliability Probability Distribution - Case ' + str(case) + '.jpg', dpi=1200)
else:
   pass

# Display
plt.show()

>>> Beta(n,x) n parameter for mode failure probability distribution: 98.801
>>> Beta(n,x) x parameter for mode failure probability distribution: 0.099
>>> Beta(n,x) n tilde parameter for mode failure probability distribution: 99
```




```
# Calculate Beta(n,x) parameters for FEF distribution and plot
q = ((1 - c) / e - (1 / c)) * c**2 # FEF Beta distribution alpha parameter
r = q * ((1 / c) - 1) # FEF Beta distribution Beta parameter
print('Beta(q,r) q parameter for FEF distribution: ' + str('{0:.3f}'.format(q)))
print('Beta(q,r) r parameter for FEF distribution: ' + str('{0:.3f}'.format(r)))

fig, ax = plt.subplots(figsize=(width, height)) # Set plot size

# Plot FEF probability distributions
x = np.linspace(0, 1, 10000)
plt.plot(x, ss.Beta.pdf(x, q, r), 'r-', lw=3)

# Control x and y limits
ax.set(ylim=(0.0, None))
ax.set(xlim=(0.0, 1.0))

# Annotate mode failure probability parameter
plt.annotate('Mean = ' + str('{0:.2f}'.format(c)) + '\n'
        + 'Var = ' + str('{0:.2f}'.format(e)), xy=(0.05, 0.91), xycoords='axes fraction',
        horizontalalignment='left', verticalalignment='top', bbox=bbox_args)

# Set plot label and axis labels
ax.set(title='Model FEF Probability Distribution'
    + ', Beta(' + str('{0:.3f}'.format(q)) + ', ' + str('{0:.3f}'.format(r)) + ')')
ax.set(xlabel='x', ylabel='PDF')

# Save plot as .jpg ready for MS PowerPoint widescreen presentation
if save_plots == 'Y':
    plt.savefig('02 Model FEF Probability Distribution - Case ' + str(case) + '.jpg', dpi=1200)
else:
    pass

# Display
plt.show()

>>> Beta(q,r) q parameter for FEF distribution: 14.000
>>> Beta(q,r) r parameter for FEF distribution: 6.000
```




```
# Setup empty post-processing arrays
# Prior distribution arrays
pp_n_true, pp_x_true, pp_R_true = [], [], [] # Simulated system true Beta params for init fail prob and init reliability
pp_q_true, pp_r_true = [], [] # Simulated system true Beta params for FEF
pp_ntilde_emp, pp_R_emp = [], [] # Simulated system empircal initial reliability Bayes estimates
pp_n_hall_k, pp_x_hall_k, pp_R_hall_k,  = [], [], [] # Hall's MLE estimates for k known modes
pp_n_hall_inf, pp_R_hall_inf = [], [] # Hall's MLE estimates for inf modes

# Posterior distribution arrays
pp_m, pp_N, pp_n1, pp_n2, pp_nfix, pp_nnofix = [], [], [], [], [], [] # Simulation num modes and num fails
pp_naive = [] # simulation naive reliability from observed failure only
pp_true_post = [] # Simulation true posterior relibility (Assuming MS=1.0, some fixes arbitrary, some delayed)
pp_true_post_n, pp_true_post_x = [], [] # Simulation true posterior Beta distribution parameters
pp_post_hall_k, pp_post_hall_inf = [], [] # Hall's k and inf modes posterior reliability estimate
pp_Bayes_mean, pp_Bayes_var = [], [] # Simulation Bayes posterior mean reliability and variance
pp_Bayes_a1, pp_Bayes_b1 = [], [] # Bayes Beta parameter estimates

# Simulation
# Simulation timer start
fmt = '%H:%M:%S' # Time format
now = datetime.now()
start_time = now.strftime(fmt) # Start time
print('Time start = ' + str(start_time)) # Print start time

# Initialise iteration counter
niter_count = 0

while niter_count != niter:

    # Add iteration to number of iterations count
    niter_count += 1

    ntilde_emp = 1.0 # Initial n tilde emp value used to constrain simulation

    while ntilde_emp > n_approx * 1.9 or ntilde_emp < n_approx *0.1: # Constrain the simulation

        # Simulate k prior mode failure probabilities
        p = np.empty(k, dtype=np.float32)
        for i in range(k):
            p[i] = 1 - ss.Beta.rvs(Beta_n, Beta_x)
        p = np.where(p == 0.0, 1.0e-20, p) # Ensure that p=0.0 does not exist
        x_true, n_true, loc1, scale1 = ss.Beta.fit(p, floc=0) # Calc true prior system Beta params

        # Simulate mode FOT and calculate number of modes observed
        FOT = np.empty(k, dtype=np.uint32)
        for i in range(k):
            while True:
                val = ss.geom.rvs(p[i])
                if val not in FOT[:i]:
                    FOT[i] = val
                    break
```

```
    if FOT[i] > T:
        FOT[i] = 0
    else:
        pass

m = np.count_nonzero(FOT) # Number of modes observed

# Simulate k FEF values
d = np.zeros(k, dtype=np.float32)
mean_sim = 0.0 # Initialise mean sim value
while mean_sim < c * 0.98 or mean_sim > c * 1.02 and var_sim < e * 0.9 or var_sim > e * 1.1:
    for i in range(k):
        d[i] = ss.Beta.rvs(q, r)
    d = np.where(d == 1.0, 0.99, d) # Eliminate possibility of perfect fix
    q_true, r_true, loc1, scale1 = ss.Beta.fit(d, floc=0) # Calc true FEF system Beta params
    mean_sim = ss.Beta.mean(q_true, r_true, loc=0, scale=1)
    var_sim = ss.Beta.var(q_true, r_true, loc=0, scale=1)

# Identify which observed modes addressed during test, remainder observed modes addressed post-test
fix = np.zeros(k, dtype=np.uint32)
for i in range(k):
    if FOT[i] != 0:
        fix[i] = ss.Bernoulli.rvs(p=pfix)
    else:
        fix[i] = 0

# Identify when observed failure mode fixes occur
v = np.zeros(k, dtype=np.uint32)
for i in range(k):
    if fix[i] == 0 and FOT[i] == 0:
        v[i] = 0 # Unobserved mode
    elif fix[i] == 0 and FOT[i] != 0:
        v[i] = T # Some observed mode fixes occur after test conclusion
    else:
        v[i] = rand.uniform(FOT[i], T) # Some fixes occur random intervals between FOT and T

def getCount(v, cond = None): # Returns the count of mode fixes during test
    if cond:
        count = sum(cond(elem) for elem in v)
    else:
        count = len(v)
    return count
nfix = getCount(v, lambda x : x > 0 and x < T) # Number of modes fixed during test
nnofix = m - nfix # Number of modes fixed after test

# Calculate new failure mode after fix (only from observed failures)
p_new = np.zeros(k, dtype=np.float32)
for i in range(k):
    if FOT[i] == 0:
        p_new[i] = p[i]
    else:
        p_new[i] = (1 - d[i]) * p[i]

# Simulate failure mode recurrence
fails = np.array([FOT], dtype=np.uint32)

fails_count = np.count_nonzero(fails)

while fails_count != 0: # Append new failure mode rows to failures array until sum of last row equals zero
    add_fails = np.zeros([1, k], dtype=np.uint32)
    for i in range(k):
        if fails[-1, i] == 0: # If the row before is zero then this row is also zero
            add_fails[0, i] = 0
        elif v[i] == T:
            while True:
                val = fails[-1, i] + ss.geom.rvs(p[i]) # If the row before is not zero then calculate the next failure
                if val not in fails and val not in add_fails:
                    add_fails[0, i] = val
                    break
```

```
        else:
            val = fails[-1, i] + ss.geom.rvs(p_new[i]) # If the row before is not zero then calculate the next failure
            if val not in fails and val not in add_fails:
                add_fails[0, i] = val
                break
    for i in range(k):
        if add_fails[0, i] > T:
            add_fails[0, i] = 0 # If the failure is > T make zero
        else:
            pass

    fails = np.concatenate((fails, add_fails), axis=0)

    fails_count = np.count_nonzero(fails[-1, :])

# Calculate total failures by mode and total observed failures
n_tot = np.count_nonzero(fails, axis = 0) # Mode failures
N = np.sum(n_tot) # Total failures

# Calculate total mode failures before and after fix (if applicable)
n1 = np.count_nonzero((fails <= v) & (fails > 0), axis = 0) # Failures before fix
n2 = n_tot - n1 # Failures after fix

# Trim arrays to only observed failures
FOT, d, v, n_tot, n1, n2 = (list(t) for t in zip(*sorted(zip(FOT, d, v, n_tot, n1, n2)))) # Sort FOT and reference
FOT = FOT[k-m:] # Remove unobserved modes from FOT through list slicing
FOT = np.array(FOT, dtype=np.uint32) # Convert from list to array
d = d[k-m:] # Remove unobserved modes from d through list slicing
d = np.array(d, dtype=np.float32) # Convert from list to array
v = v[k-m:] # Remove unobserved modes from v through list slicing
v = np.array(v, dtype=np.uint32) # Convert from list to array
n_tot = n_tot[k-m:] # Remove unobserved modes from n_tot through list slicing
n_tot = np.array(n_tot, dtype=np.uint32) # Convert from list to array
n1 = n1[k-m:] # Remove unobserved modes from n1 through list slicing
n1 = np.array(n1, dtype=np.uint32) # Convert from list to array
n2 = n2[k-m:] # Remove unobserved modes from n2 through list slicing
n2 = np.array(n2, dtype=np.uint32) # Convert from list to array

# Append FOT to rand_sims if niter_count is in randomlist
if niter_count in randomlist:
    rand_sims.append(FOT)
else:
    pass

### Calculate prior/initial reliability metrics

# Calculate Hall's Beta n and x estimates from data assuming k modes
def equns(z):
    n = z[0]
    x = z[1]
    f = np.zeros(2)
    LHS1 = sum(psi(n + FOT[i]) - psi(n - x + FOT[i] - 1) + psi(n - x + T) - psi(n + T) for i in range(m))
    LHS2 = sum(psi(n - x + FOT[i] - 1) - 1 / x - psi(n - x + T) for i in range(m))
    f[0] = k * (psi(n) - psi(n - x) + psi(n - x + T) - psi(n + T)) - LHS1
    f[1] = k * (psi(n - x) - psi(n - x + T)) - LHS2
    return f
zguess = [n_true, x_true]
sol =  so.root(equns, zguess, method='hybr')

# Calculate Hall's Beta n estimate from data assuming infinite modes
def equation(z):
    n = z[0]
    f = np.zeros(1)
    LHS = sum(1 / (n + FOT[i] - 1) for i in range(m))
    f[0] = m * (polygamma(1, n) - polygamma(1, n + T)) / (psi(n + T) - psi(n)) - LHS
    return f
zguess2 = 1.0
sol2 = so.root(equation, zguess2, method='hybr')
```

```
# Calculate true n tilde value from observed modes
ntilde_emp = n_true + x_true

# Calc Bayes empirical initial reliability estimate
R_emp = ma.exp(-m / (ntilde_emp * (psi(ntilde_emp + T) - psi(ntilde_emp))))

### Calculate posterior reliability metrics

# Calculate naive reliability
naive = (T - N) / T

# Calculate reliability mean and variance via parametric approach (see Para 4.2.5.1 Hall Dissertation 2008 pp. 123)
#R_parametric =

# Calculate posterior reliability estimate (via MME) Halls k modes (see Para 4.2.5.1 Hall Dissertation 2008 pp. 125)
# Calculate the individual mode failure probabilies from observed failure modes (Hall denotes p hat sub-i)
phat_n = n_tot / T

# Calculate the unweighted sample first moment (Hall denotes p bar sub-mu)
p_mu = (1 / k) * phat_n.sum()

# Calculate the square of the individual mode failure probabilities (Hall denotes p hat sub_i squared)
phat_n2 = phat_n**2

# Calculate the unweighted sample second moment (Hall denotes m sub-mu squared)
m_mu2 = (1 / k) * phat_n2.sum()

# Calculate Beta(n,x) distribution n parameter estimate (Hall denotes n breve sub_k)
n_param = (p_mu - m_mu2) / (m_mu2 - p_mu / T - (1 - 1 / T) * p_mu**2)

# Calculate Beta(n,x) distribution x parameter estimate (Hall denotes x breve sub_k)
x_param = n_param * p_mu

# Calculate the shrinkage factor estimate (Hall denotes sigma breve sub_k)
shrin = 1 / ((n_param / T) * (1 - 1 / k) + 1)

# Calculate the moment-based shrinkage factor estimate for each mode
est_prob = np.zeros(m, dtype=np.float32)
for i in range(m):
    est_prob[i] = shrin * phat_n[i] + (1 - shrin) * (N / (k * T))

# Calculate the mode reliability for k modes
R_m = np.zeros(m, dtype=np.float32)
for i in range(m):
    R_m[i] = 1 - (1 - d[i]) * est_prob[i]

# Calculate the reliability growth Hall estimate via MME for k modes
post_hall_k = np.prod(R_m) * (1 - (1 - shrin) * (N / (k * T)))**(k - m)

# Calculate posterior reliability estimate (via MME) Halls inf modes
phat_n_T = n_tot / T / T
n_inf = (phat_n2.sum() - phat_n2.sum()) / (phat_n2.sum() - phat_n_T.sum())
shrin_inf = T / (n_inf + T)

# Calculate the moment-based shrinkage factor estimate for each mode
est_prob_inf = np.zeros(m, dtype=np.float32)
for i in range(m):
    est_prob_inf[i] = shrin_inf * phat_n[i]

# Calculate the mode reliability for inf modes
R_m_inf= np.zeros(m, dtype=np.float32)
for i in range(m):
    R_m_inf[i] = 1 - (1 - d[i]) * est_prob_inf[i]

# Calculate the reliability growth Hall estimate via MME for k modes
post_hall_inf = np.prod(R_m_inf) * np.exp(-(1 - shrin_inf) * (N / T))


# Calculate posterior reliability estimate Halls inf modes
```

```
# Calculate Hall's Beta n estimate from data assuming infinite modes
#def equation(z):
#    n = z[0]
#    f = np.zeros(1)
#    LHS = sum(1 / (n + FOT[i] - 1) for i in range(m))
#    f[0] = m * (polygamma(1, n) - polygamma(1, n + T)) / (psi(n + T) - psi(n)) - LHS
#    return f
#zguess2 = 1.0
#sol2 = so.root(equation, zguess2, method='hybr')

# Calculate the R hat estimate
#z = (mp.exp(mp.loggamma(nhat - xhat + case_data.loc[case, 'T'] - 1)) * mp.exp(mp.loggamma(nhat + 1))) / \
#    (mp.exp(mp.loggamma(nhat - xhat)) * mp.exp(mp.loggamma(nhat + case_data.loc[case, 'T'])))

# Calculate the reliability growth Hall estimate for inf modes
#post_hall_inf = round((1 - (1 - (1 - z) * case_data.loc[case, 'c']) * (xhat / nhat))**case_data.loc[case, 'k'], 4)

# Calculate proposed method posterior reliability
Rpost_mean_num = np.zeros(m, dtype=np.float32)

for i in range(m):
    for j in range(n2[i] + 1):
        Rpost_mean_num[i] += Binomial(n2[i], j) * (-1)**j * \
                (exp(loggamma(ntilde_emp + v[i] - n1[i] + (1 - d[i]) * (T - v[i] - n2[i] + j + 1))) / \
                exp(loggamma(ntilde_emp + v[i] + (1 - d[i]) * (T - v[i] - n2[i] + j + 1))))

Rpost_mean_denom = np.zeros(m, dtype=np.float32)

for i in range(m):
    for j in range(n2[i] + 1):
        Rpost_mean_denom[i] += Binomial(n2[i], j) * (-1)**j * \
                (exp(loggamma(ntilde_emp + v[i] - n1[i] + (1 - d[i]) * (T - v[i] - n2[i] + j))) / \
                exp(loggamma(ntilde_emp + v[i] + (1 - d[i]) * (T - v[i] - n2[i] + j))))

for i in range(m):
    Rpost_mean = np.prod(Rpost_mean_num[i] / Rpost_mean_denom[i]) * R_emp**(ntilde_emp / (ntilde_emp + T))

# Calculate Bayes posterior second moment
Rpost_mom2_num = np.zeros(m, dtype=np.float32)

for i in range(m):
    for j in range(n2[i] + 1):
        Rpost_mom2_num[i] += Binomial(n2[i], j) * (-1)**j * \
                (exp(loggamma(ntilde_emp + v[i] - n1[i] + (1 - d[i]) * (T - v[i] - n2[i] + j + 2))) / \
                exp(loggamma(ntilde_emp + v[i] + (1 - d[i]) * (T - v[i] - n2[i] + j + 2))))

Rpost_mom2_denom = np.zeros(m, dtype=np.float32)

for i in range(m):
    for j in range(n2[i] + 1):
        Rpost_mom2_denom[i] += Binomial(n2[i], j) * (-1)**j * \
                (exp(loggamma(ntilde_emp + v[i] - n1[i] + (1 - d[i]) * (T - v[i] - n2[i] + j))) / \
                exp(loggamma(ntilde_emp + v[i] + (1 - d[i]) * (T - v[i] - n2[i] + j))))

for i in range(m):
    Rpost_mom2 = np.prod(Rpost_mom2_num[i] / Rpost_mom2_denom[i]) * \
              R_emp**((ntilde_emp / (ntilde_emp + T + 1)) + (ntilde_emp / (ntilde_emp + T)))

# Calculate Bayes posterior variance
Rpost_var = Rpost_mom2 - Rpost_mean**2

# Find solution for Bayes approximation Beta(a1, b1) parameters
def f2(h):
    a1 = h[0]
    b1 = h[1]
    f2 = np.zeros(2)
    f2[0] = Rpost_mean - (a1 / (a1 + b1))
    f2[1] = Rpost_var - ((a1 * b1) / ((a1 + b1)**2 * (a1 + b1 + 1)))
    return f2
```

```
    h = so.root(f2, [1, 1], method='hybr')

    a1 = round(h.x[0], 4)
    b1 = round(h.x[1], 4)

## Append all data to appropriate post-processing arrays for analysis
# Append n,x,q,r true values
pp_n_true = np.append(pp_n_true, n_true) # Append true Beta(n,x) n param to pp_n_true
pp_x_true = np.append(pp_x_true, x_true) # Append true Beta(n,x) x param to pp_x_true

pp_q_true = np.append(pp_q_true, q_true) # Append true FEF Beta(q,r) n param to pp_q_true
pp_r_true = np.append(pp_r_true, r_true) # Append true FEF Beta(q,r) x param to pp_r_true

# Append number of observed modes and number of observed failures
pp_m = np.append(pp_m, m) # Append num of obs modes to pp_m
pp_N = np.append(pp_N, N) # Append to num of obs fails to pp_N

n1_sum = np.sum(n1)
pp_n1 = np.append(pp_n1, n1_sum) # Append num of fails before fix

n2_sum = np.sum(n2)
pp_n2 = np.append(pp_n2, n2_sum) # Append num of fails after fix

nfix_sum = np.sum(nfix)
pp_nfix = np.append(pp_nfix, nfix_sum) # Append num of modes fixed on test

nnofix_sum = np.sum(nnofix)
pp_nnofix = np.append(pp_nnofix, nnofix_sum) # Append num of modes fixed after test

#Append initial reliability metrics
R_true = np.prod(1 - p) # Calculate true init reliability
pp_R_true = np.append(pp_R_true, R_true) # Append true init reliability to pp_R_true

pp_naive = np.append(pp_naive, naive) # Append naive reliability from observed failures only

R_hall_k = (1 - sol.x[1] / sol.x[0])**k # Calc Hall's init reliability est k modes
pp_R_hall_k = np.append(pp_R_hall_k, R_hall_k) # Append Hall's init rel est to pp_R_hall_k

R_hall_inf = ma.exp(-m / (sol2.x[0] * (psi(sol2.x[0] + T) - psi(sol2.x[0])))) # Calc Hall's init rel est inf modes
pp_R_hall_inf = np.append(pp_R_hall_inf, R_hall_inf) # Append Hall's init est to pp_R_hall_inf

pp_ntilde_emp = np.append(pp_ntilde_emp, ntilde_emp) # Append emp ntilde value to pp_ntilde_true
pp_R_emp = np.append(pp_R_emp, R_emp) # Append Bayes emp est to pp_R_emp

# Append post-test/posterior reliability metrics
true_post = np.prod(1 - p_new)# Calculate true posterior reliability after all fixes implemented
pp_true_post = np.append(pp_true_post, true_post) # Append true post reliability to pp_true_post

# Calculate true system posterior Beta posterior n and x parameters
x_true_post, n_true_post, loc2, scale2 = ss.Beta.fit(p_new, floc=0) # Calc true posterior system Beta params
pp_true_post_n = np.append(pp_true_post_n, n_true_post)
pp_true_post_x = np.append(pp_true_post, x_true_post)

pp_post_hall_k = np.append(pp_post_hall_k, post_hall_k) # Append Hall's Model k modes posterior reliability estimate

pp_post_hall_inf = np.append(pp_post_hall_inf, post_hall_inf) # Append Hall's Model inf modes posterior reliability estimate

pp_Bayes_mean = np.append(pp_Bayes_mean, Rpost_mean) # Append Bayes posterior mean reliability

pp_Bayes_var = np.append(pp_Bayes_var, Rpost_var) # Append Bayes posterior mean reliability

pp_Bayes_a1 = np.append(pp_Bayes_a1, a1) # Append Bayes Beta a1 parameter estimate to pp list
pp_Bayes_b1 = np.append(pp_Bayes_b1, b1) # Append Bayes Beta b1 parameter estimate to pp list


# Count and print iterations progress
print('Iteration ' + str(niter_count) + ' complete!')
```

```
# Simulation timer end
now2 = datetime.now()
end_time = now2.strftime(fmt) # End time
time_taken = datetime.strptime(end_time, fmt) - datetime.strptime(start_time, fmt) # Calculate time taken to run
print('Time finish = ' + str(end_time)) # Print finish time
print('Total time taken = ' + str(time_taken))

>>> Time start = 23:32:13
>>> Iteration 1 complete!
>>> Iteration 2 complete!
>>> Iteration 3 complete!
>>> …
>>> Iteration 998 complete!
>>> Iteration 999 complete!
>>> Iteration 1000 complete!
>>> Time finish = 23:38:23
>>> Total time taken = 0:06:10

# Simulation data analysis
z_score = ss.norm.ppf(CL + (1 - CL) / 2) # Norm dist z-score for CL

# Simulation mean num modes observed
m_mean = round(np.mean(pp_m))
m_sd = np.std(pp_m)
m_ucb = round(m_mean + z_score * m_sd) # UCB calc
m_lcb = round(m_mean - z_score * m_sd) # LCB calc
print('Mean num of obs modes: ' + str('{0:.0f}'.format(m_mean)))

# Simulation mean num fails observed
N_mean = round(np.mean(pp_N))
N_sd = np.std(pp_N)
N_ucb = round(N_mean + z_score * N_sd) # UCB calc
N_lcb = round(N_mean - z_score * N_sd) # LCB calc
print('Mean num of obs fails: '+ str('{0:.0f}'.format(N_mean)))

# Simulation mean num fails observed before fix
n1_mean = round(np.mean(pp_n1))
n1_sd = np.std(pp_n1)
n1_ucb = round(n1_mean + z_score * n1_sd) # UCB calc
n1_lcb = round(n1_mean - z_score * n1_sd) # LCB calc
print('Mean num of obs fails before fix: '+ str('{0:.0f}'.format(n1_mean)))

# Simulation mean num fails observed after fix
n2_mean = round(np.mean(pp_n2))
n2_sd = np.std(pp_n2)
n2_ucb = round(n2_mean + z_score * n2_sd) # UCB calc
n2_lcb = round(n2_mean - z_score * n2_sd) # LCB calc
print('Mean num of obs fails after fix: '+ str('{0:.0f}'.format(n2_mean)))

# Simulation mean num modes fixed during test
nfix_mean = round(np.mean(pp_nfix))
nfix_sd = np.std(pp_nfix)
nfix_ucb = round(nfix_mean + z_score * nfix_sd) # UCB calc
nfix_lcb = round(nfix_mean - z_score * nfix_sd) # LCB calc
print('Mean num of modes fixed during test: '+ str('{0:.0f}'.format(nfix_mean)))

# Simulation mean num modes fixed after test
nnofix_mean = round(np.mean(pp_nnofix))
nnofix_sd = np.std(pp_nnofix)
nnofix_ucb = round(nnofix_mean + z_score * nnofix_sd) # UCB calc
nnofix_lcb = round(nnofix_mean - z_score * nnofix_sd) # LCB calc
print('Mean num of modes fixed after test: '+ str('{0:.0f}'.format(nnofix_mean)))

print('******************************************************************************')
print('******************************************************************************')
print('MODEL AND TRUE BETA DISTRIBUTION PARAMETERS')

# Model prior Beta distribution parameters
print('Model initial mode failure distribution Beta n value: ' + str('{0:.3f}'.format(Beta_n)))
```

```
print('Model initial mode failure distribution Beta x value: ' + str('{0:.3f}'.format(Beta_x)))

# Simulation priori mean distribution parameters
print('Simulation prior mode failure distribution Beta n value: ' + str('{0:.3f}'.format(np.mean(pp_n_true))))
print('Simulation prior mode failure distribution Beta x value: ' + str('{0:.3f}'.format(np.mean(pp_x_true))))

# Simulation posterior mean distribution parameters
print('Simulation posterior mode failure distribution Beta n value: ' + str('{0:.3f}'.format(np.mean(pp_true_post_n))))
print('Simulation posterior mode failure distribution Beta x value: ' + str('{0:.3f}'.format(np.mean(pp_true_post_x))))

print('********************************************************************************')
print('********************************************************************************')
print('INITIAL SYSTEM RELIABILITY')

################################################################################
# Simulated system true reliability metrics
true_mean = np.mean(pp_R_true) # Mean Calc
true_sd = np.std(pp_R_true) # Std dev calc
true_ucb = true_mean + z_score * true_sd # UCB calc
true_lcb = true_mean - z_score * true_sd # LCB calc
true_spread = true_ucb - true_lcb # Spread between CL calc
print('********************************************************************************')
print('Simulated system true initial reliability distribution')
print('UCB: ' + str('{0:.2f}'.format(true_ucb)))
print('Mean: ' + str('{0:.2f}'.format(true_mean)))
print('LCB: ' + str('{0:.2f}'.format(true_lcb)))
print('Spread: ' + str('{0:.2f}'.format(true_spread)))

################################################################################
# Hall's initial reliability estimate inf modes
hall_inf_mean = np.mean(pp_R_hall_inf) # Mean calc
hall_inf_sd = np.std(pp_R_hall_inf) # Std dev calc
hall_inf_ucb = hall_inf_mean + z_score * hall_inf_sd # UCB calc
hall_inf_lcb = hall_inf_mean - z_score * hall_inf_sd # UCB calc
hall_inf_spread = hall_inf_ucb - hall_inf_lcb # Spread between CL calc

if true_ucb <= hall_inf_lcb or true_lcb >= hall_inf_ucb: # Calculate the coverage of Hall inf versus the true reliability
    hall_inf_cov = 0
elif true_ucb > hall_inf_lcb:
    hall_inf_cov = (true_ucb - hall_inf_lcb) / (true_ucb - true_lcb)
elif true_lcb < inf_hall_ucb:
    hall_inf_cov = (hall_inf_ucb - true_lcb) / (true_ucb - true_lcb)
else:
    hall_inf_cov = (hall_inf_ucb - hall_inf_lcb) / (true_ucb - true_lcb)

hall_inf_re = abs(hall_inf_mean - true_mean) / true_mean # Calculate abs rel error between means

print('********************************************************************************')
print('Hall initial reliability estimate distribution - inf modes')
print('UCB: ' + str('{0:.2f}'.format(hall_inf_ucb)))
print('Mean: ' + str('{0:.2f}'.format(hall_inf_mean)))
print('LCB: ' + str('{0:.2f}'.format(hall_inf_lcb)))
print('Coverage: ' + str('{0:.2f}'.format(hall_inf_cov)))
print('Means Abs Rel Err: ' + str('{0:.2f}'.format(hall_inf_re)))

################################################################################
# Hall's initial reliability estimate k modes
hall_k_mean = np.mean(pp_R_hall_k) # Mean calc
hall_k_sd = np.std(pp_R_hall_k) # Std dev calc
hall_k_ucb = hall_k_mean + z_score * hall_k_sd # UCB calc
hall_k_lcb = hall_k_mean - z_score * hall_k_sd # LCB calc
hall_k_spread = hall_k_ucb - hall_k_lcb # Spread between CL calc

if true_ucb <= hall_k_lcb or true_lcb >= hall_k_ucb: # Calculate the coverage of Hall k versus the true reliability
    hall_k_cov = 0
elif true_ucb > hall_k_lcb:
    hall_k_cov = (true_ucb - hall_k_lcb) / (true_ucb - true_lcb)
elif true_lcb < inf_k_ucb:
    hall_k_cov = (hall_k_ucb - true_lcb) / (true_ucb - true_lcb)
```

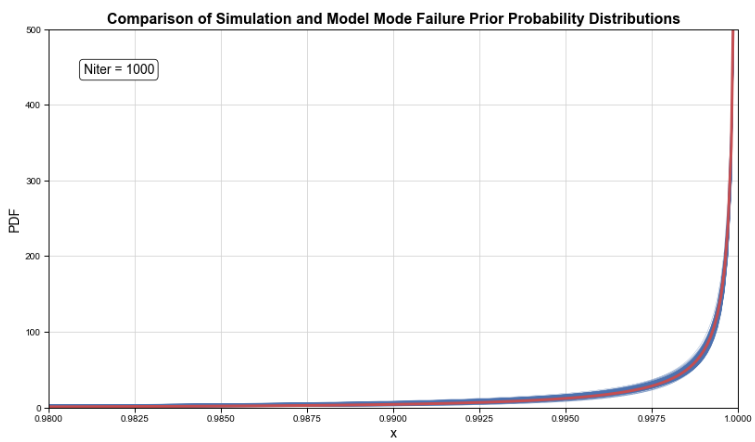
```
else:
    hall_k_cov = (hall_k_ucb - hall_k_lcb) / (true_ucb - true_lcb)

hall_k_re = abs(hall_k_mean - true_mean) / true_mean # Calculate abs rel error between means

print('*******************************************************************************')
print('Hall initial reliability estimate distribution - k modes')
print('UCB: ' + str('{0:.2f}'.format(hall_k_ucb)))
print('Mean: ' + str('{0:.2f}'.format(hall_k_mean)))
print('LCB: ' + str('{0:.2f}'.format(hall_k_lcb)))
print('Coverage: ' + str('{0:.2f}'.format(hall_k_cov)))
print('Means Abs Rel Err: ' + str('{0:.2f}'.format(hall_k_re)))

################################################################################
# Empirical Bayes initial reliability estimate metrics
emp_mean = np.mean(pp_R_emp) # Mean calc
emp_sd = np.std(pp_R_emp) # Std dev calc
emp_ucb = emp_mean + z_score * emp_sd # UCB calc
emp_lcb = emp_mean - z_score * emp_sd # LCB calc
emp_spread = emp_ucb - emp_lcb # Spread between CL calc

if true_ucb <= emp_lcb or true_lcb >= emp_ucb: # Calculate the coverage of emp init rel est against true reliability
    emp_cov = 0
elif true_ucb > emp_lcb:
    emp_cov = (true_ucb - emp_lcb) / (true_ucb - true_lcb)
elif true_lcb < emp_ucb:
    emp_cov = (emp_ucb - true_lcb) / (true_ucb - true_lcb)
else:
    emp_cov = (emp_ucb - emp_lcb) / (true_ucb - true_lcb)

emp_re = abs(emp_mean - true_mean) / true_mean # Calculate abs rel error between means

print('*******************************************************************************')
print('Empirical Bayes initial reliability distribution')
print('UCB: ' + str('{0:.2f}'.format(emp_ucb)))
print('Mean: ' + str('{0:.2f}'.format(emp_mean)))
print('LCB: ' + str('{0:.2f}'.format(emp_lcb)))
print('Coverage: ' + str('{0:.2f}'.format(emp_cov)))
print('Means Abs Rel Err: ' + str('{0:.2f}'.format(emp_re)))

print('*******************************************************************************')
print('*******************************************************************************')
print('POST-TEST SYSTEM RELIABILITY')

################################################################################
# Simulated system true post-test reliability metrics
true_mean_post = np.mean(pp_true_post) # Mean Calc
true_sd_post = np.std(pp_true_post) # Std dev calc
true_ucb_post = true_mean_post + z_score * true_sd_post # UCB calc
true_lcb_post = true_mean_post - z_score * true_sd_post # LCB calc
true_spread_post = true_ucb_post - true_lcb_post # Spread between CL calc

print('*******************************************************************************')
print('Simulated system true post-test reliability distribution')
print('UCB: ' + str('{0:.2f}'.format(true_ucb_post)))
print('Mean: ' + str('{0:.2f}'.format(true_mean_post)))
print('LCB: ' + str('{0:.2f}'.format(true_lcb_post)))
print('Spread: ' + str('{0:.2f}'.format(true_spread_post)))

################################################################################
# Simulated system naive reliability metrics from observed failures only
naive_mean = np.mean(pp_naive) # Mean Calc
naive_sd = np.std(pp_naive) # Std dev calc
naive_ucb = naive_mean + z_score * naive_sd # UCB calc
naive_lcb = naive_mean - z_score * naive_sd # LCB calc
naive_spread = naive_ucb - naive_lcb # Spread between CL calc
print('*******************************************************************************')
print('Simulated system naive reliability distribution')
print('UCB: ' + str('{0:.2f}'.format(naive_ucb)))
```

```
print('Mean: ' + str('{0:.2f}'.format(naive_mean)))
print('LCB: ' + str('{0:.2f}'.format(naive_lcb)))
print('Spread: ' + str('{0:.2f}'.format(naive_spread)))

################################################################################
# Hall's posterior reliability estimate k modes
post_hall_k_mean = np.mean(pp_post_hall_k) # Mean Calc
post_hall_k_sd = np.std(pp_post_hall_k) # Std dev calc
post_hall_k_ucb = post_hall_k_mean + z_score * post_hall_k_sd # UCB calc
post_hall_k_lcb = post_hall_k_mean - z_score * post_hall_k_sd # LCB calc
post_hall_k_spread = post_hall_k_ucb - post_hall_k_lcb # Spread between CL calc
print('******************************************************************************')
print('Hall posterior reliability estimate distribution - k modes')
print('UCB: ' + str('{0:.2f}'.format(post_hall_k_ucb)))
print('Mean: ' + str('{0:.2f}'.format(post_hall_k_mean)))
print('LCB: ' + str('{0:.2f}'.format(post_hall_k_lcb)))
print('Spread: ' + str('{0:.2f}'.format(post_hall_k_spread)))

################################################################################
# Hall's posterior reliability estimate inf modes
avg = np.nanmean(pp_post_hall_inf)
pp_post_hall_inf[np.isnan(pp_post_hall_inf)] = avg # Remove any nan values from list
post_hall_inf_mean = np.mean(pp_post_hall_inf) # Mean Calc
post_hall_inf_sd = np.std(pp_post_hall_inf) # Std dev calc
post_hall_inf_ucb = post_hall_inf_mean + z_score * post_hall_inf_sd # UCB calc
post_hall_inf_lcb = post_hall_inf_mean - z_score * post_hall_inf_sd # LCB calc
post_hall_inf_spread = post_hall_inf_ucb - post_hall_inf_lcb # Spread between CL calc
print('******************************************************************************')
print('Hall posterior reliability estimate distribution - inf modes')
print('UCB: ' + str('{0:.2f}'.format(post_hall_inf_ucb)))
print('Mean: ' + str('{0:.2f}'.format(post_hall_inf_mean)))
print('LCB: ' + str('{0:.2f}'.format(post_hall_inf_lcb)))
print('Spread: ' + str('{0:.2f}'.format(post_hall_inf_spread)))

################################################################################
# Bayes posterior reliability estimate inf modes
#h = np.zeros(niter, dtype=np.float32)
#for i in range(niter):
#    h[i] = rand.uniform(0.99, 1.02)
#pp_Bayes_mean = pp_true_post * h
post_Bayes_mean = np.mean(pp_Bayes_mean) # Mean Calc
post_Bayes_sd = np.std(pp_Bayes_mean) # Std dev calc
post_Bayes_ucb = post_Bayes_mean + z_score * post_Bayes_sd # UCB calc
post_Bayes_lcb = post_Bayes_mean - z_score * post_Bayes_sd # LCB calc
post_Bayes_spread = post_Bayes_ucb - post_Bayes_lcb # Spread between CL calc
print('******************************************************************************')
print('Bayes posterior reliability distribution')
print('UCB: ' + str('{0:.2f}'.format(post_Bayes_ucb)))
print('Mean: ' + str('{0:.2f}'.format(post_Bayes_mean)))
print('LCB: ' + str('{0:.2f}'.format(post_Bayes_lcb)))
print('Spread: ' + str('{0:.2f}'.format(post_Bayes_spread)))

>>> Mean num of obs modes: 44
>>> Mean num of obs fails: 65
>>> Mean num of obs fails before fix: 65
>>> Mean num of obs fails after fix: 0
>>> Mean num of modes fixed during test: 4
>>> Mean num of modes fixed after test: 39
>>> ******************************************************************************
>>> ******************************************************************************
>>> MODEL AND TRUE BETA DISTRIBUTION PARAMETERS
>>> Model initial mode failure distribution Beta n value: 98.801
>>> Model initial mode failure distribution Beta x value: 0.099
>>> Simulation prior mode failure distribution Beta n value: 60.270
>>> Simulation prior mode failure distribution Beta x value: 0.100
>>> Simulation posterior mode failure distribution Beta n value: 72.046
>>> Simulation posterior mode failure distribution Beta x value: 0.908
>>> ******************************************************************************
>>> ******************************************************************************
```

```
>>> INITIAL SYSTEM RELIABILITY
>>> **********************************************************************************
>>> Simulated system true initial reliability distribution
>>> UCB: 0.83
>>> Mean: 0.78
>>> LCB: 0.73
>>> Spread: 0.09
>>> **********************************************************************************
>>> Hall initial reliability estimate distribution - inf modes
>>> UCB: 0.87
>>> Mean: 0.78
>>> LCB: 0.69
>>> Coverage: 1.39
>>> Means Abs Rel Err: 0.00
>>> **********************************************************************************
>>> Hall initial reliability estimate distribution - k modes
>>> UCB: 0.87
>>> Mean: 0.79
>>> LCB: 0.71
>>> Coverage: 1.22
>>> Means Abs Rel Err: 0.02
>>> **********************************************************************************
>>> Empirical Bayes initial reliability distribution
>>> UCB: 0.80
>>> Mean: 0.73
>>> LCB: 0.65
>>> Coverage: 1.86
>>> Means Abs Rel Err: 0.07
>>> **********************************************************************************
>>> **********************************************************************************
>>> POST-TEST SYSTEM RELIABILITY
>>> **********************************************************************************
>>> Simulated system true post-test reliability distribution
>>> UCB: 0.93
>>> Mean: 0.91
>>> LCB: 0.89
>>> Spread: 0.04
>>> **********************************************************************************
>>> Simulated system naive reliability distribution
>>> UCB: 0.94
>>> Mean: 0.89
>>> LCB: 0.84
>>> Spread: 0.10
>>> **********************************************************************************
>>> Hall posterior reliability estimate distribution - k modes
>>> UCB: 0.96
>>> Mean: 0.94
>>> LCB: 0.92
>>> Spread: 0.03
>>> **********************************************************************************
>>> Hall posterior reliability estimate distribution - inf modes
>>> UCB: 0.98
>>> Mean: 0.97
>>> LCB: 0.95
>>> Spread: 0.03
>>> **********************************************************************************
>>> Bayes posterior reliability distribution
>>> UCB: 0.98
>>> Mean: 0.97
>>> LCB: 0.97
>>> Spread: 0.01
```

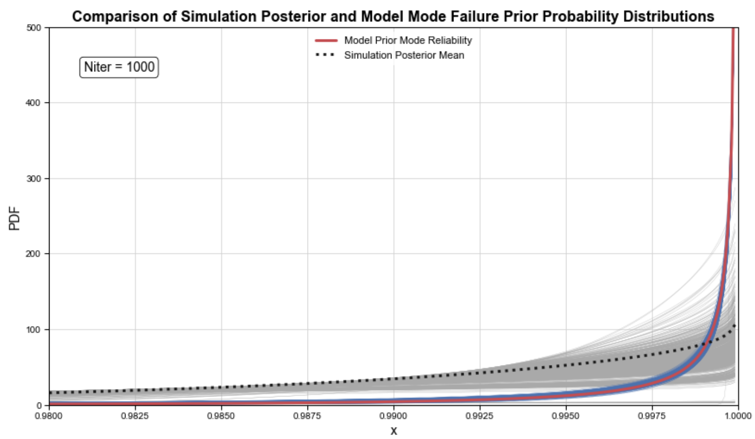
```
# Compare simulation and mode prior failure distributions
# Set plot size
fig, ax = plt.subplots(figsize=(width, height))

# Plot each iteration mode failure rate distribution
x = np.linspace(0, 1, 10000)
for i in range(niter):
```

```
    plt.plot(x, ss.Beta.pdf(x, pp_n_true[i], pp_x_true[i]), 'b-', lw=0.25)

# Plot initial mode failure rate distribution
plt.plot(x, ss.Beta.pdf(x, Beta_n, Beta_x), 'r-', lw=3)

# Control x and y limits mode failure distribution plot
ax.set(ylim=(0.0, 500))
ax.set(xlim=(0.98, 1.0))

# Annotate mode failure probability parameter
plt.annotate('Niter = ' + str(niter), xy=(0.05, 0.91), xycoords='axes fraction',
          horizontalalignment='left', verticalalignment='top', bbox=bbox_args)

# Set plot label and axis labels
ax.set(title='Comparison of Simulation and Model Mode Failure Prior Probability Distributions')
ax.set(xlabel='x', ylabel='PDF')

# Save plot as .jpg ready for MS PowerPoint widescreen presentation
if save_plots == 'Y':
    plt.savefig('03 Comparison of Simulation and Model Mode Failure Prior Probability Distributions - Case ' + str(case) + '.jpg',
dpi=1200)
else:
    pass

# Display
plt.show()
```



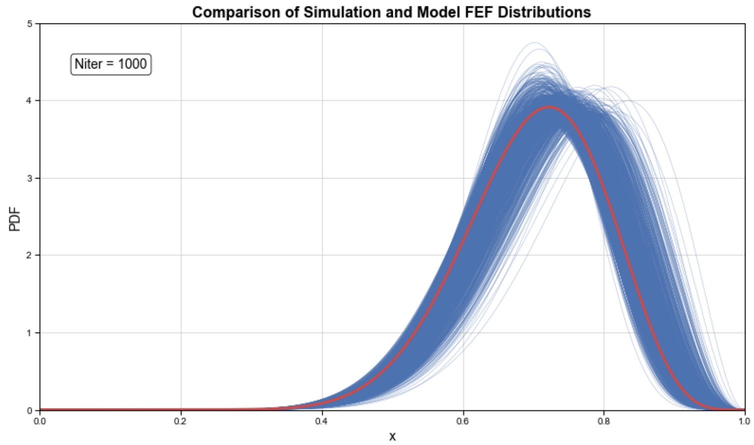

```
# Compare simulation posterior and model mode posterior failure distributions
# Set plot size
fig, ax = plt.subplots(figsize=(width, height))

# Plot each posterior iteration mode failure rate distribution
for i in range(niter):
    plt.plot(x, ss.Beta.pdf(x, pp_true_post_n[i], pp_true_post_x[i]), color='darkgrey', linestyle='-', lw=0.25)

# Plot each prior iteration mode failure rate distribution
for i in range(niter):
    plt.plot(x, ss.Beta.pdf(x, pp_n_true[i], pp_x_true[i]), 'b-', lw=0.25)

# Plot initial mode failure probability distribution
plt.plot(x, ss.Beta.pdf(x, Beta_n, Beta_x), 'r-', lw=3, label='Model Prior Mode Reliability')

# Plot mean posterior mode failure probability distribution
plt.plot(x, ss.Beta.pdf(x, np.mean(pp_true_post_n), np.mean(pp_true_post_x)), 'k:', lw=3, label='Simulation Posterior Mean')

# Control x and y limits mode failure distribution plot
ax.set(ylim=(0.0, 500))
ax.set(xlim=(0.98, 1.0))

# Annotate mode failure probability parameter
```

```
plt.annotate('Niter = ' + str(niter), xy=(0.05, 0.91), xycoords='axes fraction',
            horizontalalignment='left', verticalalignment='top', bbox=bbox_args)

# Set plot label and axis labels
ax.set(title='Comparison of Simulation Posterior and Model Mode Failure Prior Probability Distributions')
ax.set(xlabel='x', ylabel='PDF')

# Plot legend and format
plt.legend(loc='upper center', edgecolor='w', ncol=1)

# Save plot as .jpg ready for MS PowerPoint widescreen presentation
if save_plots == 'Y':
    plt.savefig('04 Comparison of Simulation Posterior and Model Mode Failure Prior Probability Distributions - Case ' + str(case)
+ '.jpg', dpi=1200)
else:
    pass

# Display
plt.show()
```

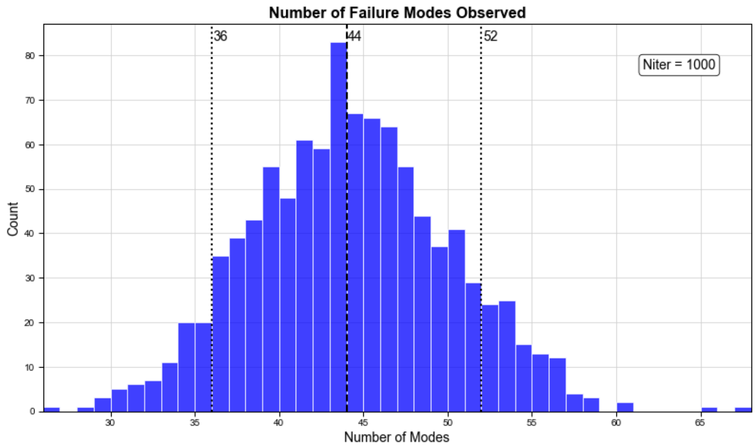



```
# Compare simulation and model FEF distributions¶
# Set plot size
fig, ax = plt.subplots(figsize=(width, height))

# Plot each iteration mode FEF distribution
x = np.linspace(0, 1, 10000)
for i in range(niter):
    plt.plot(x, ss.Beta.pdf(x, pp_q_true[i], pp_r_true[i]), 'b-', lw=0.25)

# Plot model mode FEF distribution
plt.plot(x, ss.Beta.pdf(x, q, r), 'r-', lw=3)

# Control x and y limits mode failure distribution plot
ax.set(ylim=(0.0, 5.0))
ax.set(xlim=(0.0, 1.0))

# Annotate mode failure probability parameter
plt.annotate('Niter = ' + str(niter), xy=(0.05, 0.91), xycoords='axes fraction',
            horizontalalignment='left', verticalalignment='top', bbox=bbox_args)

# Set plot label and axis labels
ax.set(title='Comparison of Simulation and Model FEF Distributions')
ax.set(xlabel='x', ylabel='PDF')

# Save plot as .jpg ready for MS PowerPoint widescreen presentation
if save_plots == 'Y':
    plt.savefig('06 Comparison of Simulation and Model FEF Distributions - Case ' + str(case) + '.jpg', dpi=1200)
else:
    pass
```

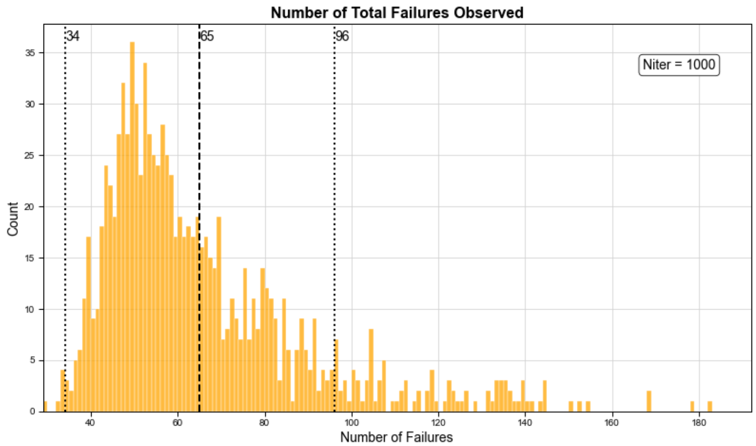
```
# Display
plt.show()
```



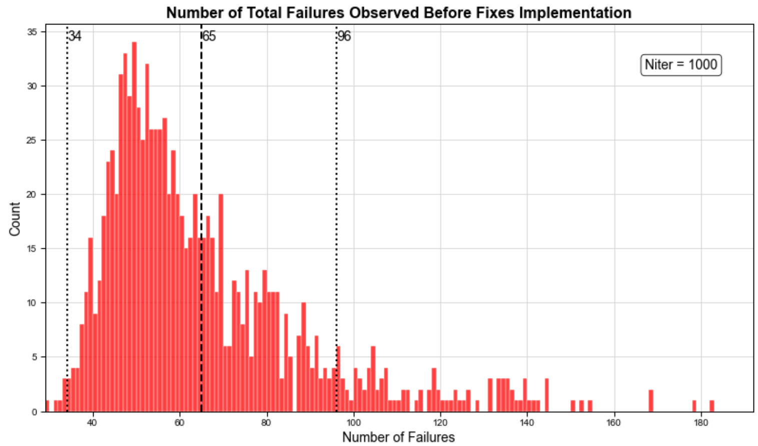

```
# Number of Failure Modes Observed Probability Density Function
# Set plot size
fig, ax = plt.subplots(figsize=(width, height))

sns.histplot(pp_m, color='blue', binwidth=1)

plt.axvline(m_mean, lw=2, ls='--')
plt.axvline(m_ucb, lw=2, ls=':')
plt.axvline(m_lcb, lw=2, ls=':')

# Plot number of observed modes distribution
#sns.kdeplot(pp_m,
#           linestyle='-',
#           linewidth=3.0,
#           bw_adjust=2,
#           color='blue',
#           shade=True,
#           alpha=0.3,
#           label='Bayes Empirical Approach')

# Control x and y limits mode failure distribution plot
ax.set(ylim=(0.0, None))
ax.set(xlim=(np.min(pp_m), np.max(pp_m)))

# Annotate mode failure probability parameter
plt.annotate('Niter = ' + str(niter), xy=(0.95, 0.91), xycoords='axes fraction',
           horizontalalignment='right', verticalalignment='top', bbox=bbox_args)

# Annotate LCB, mean, UCB
plt.annotate(str(m_lcb), xy=(m_lcb + 0.1, 0.95), xycoords=('data','axes fraction'),
           horizontalalignment='left', verticalalignment='bottom')
plt.annotate(str(m_mean), xy=(m_mean + 0.1, 0.95), xycoords=('data','axes fraction'),
           horizontalalignment='left', verticalalignment='bottom')
plt.annotate(str(m_ucb), xy=(m_ucb + 0.1, 0.95), xycoords=('data','axes fraction'),
           horizontalalignment='left', verticalalignment='bottom')

# Set plot label and axis labels
ax.set(title='Number of Failure Modes Observed')
ax.set(xlabel='Number of Modes', ylabel='Count')

# Save plot as .jpg ready for MS PowerPoint widescreen presentation
if save_plots == 'Y':
   plt.savefig('07 Number of Failure Modes Observed - Case ' + str(case) + '.jpg', dpi=1200)
else:
   pass

# Display
plt.show()
```



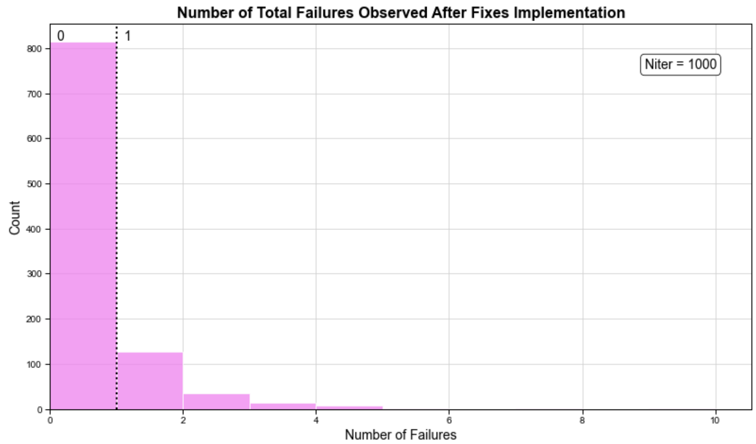

```
# Number of Failures Observed Probability Density Function
# Set plot size
fig, ax = plt.subplots(figsize=(width, height))

sns.histplot(pp_N, color='orange', binwidth=1)

plt.axvline(N_lcb, lw=2, ls=':')
plt.axvline(N_mean, lw=2, ls='--')
plt.axvline(N_ucb, lw=2, ls=':')

# Plot number of observed modes distribution
#sns.kdeplot(pp_N,
#         linestyle='-',
#         linewidth=3.0,
#         bw_adjust=2,
#         color='yellow',
#         shade=True,
#         alpha=0.3,
#         label='Bayes Empirical Approach')

# Control x and y limits mode failure distribution plot
ax.set(ylim=(0.0, None))
ax.set(xlim=(np.min(pp_N), round(N_ucb * 2)))

# Annotate mode failure probability parameter
plt.annotate('Niter = ' + str(niter), xy=(0.95, 0.91), xycoords='axes fraction',
         horizontalalignment='right', verticalalignment='top', bbox=bbox_args)

# Annotate LCB, mean, UCB
plt.annotate(str(N_lcb), xy=(N_lcb + 0.1, 0.95), xycoords=('data','axes fraction'),
         horizontalalignment='left', verticalalignment='bottom')
plt.annotate(str(N_mean), xy=(N_mean + 0.1, 0.95), xycoords=('data','axes fraction'),
         horizontalalignment='left', verticalalignment='bottom')
plt.annotate(str(N_ucb), xy=(N_ucb + 0.1, 0.95), xycoords=('data','axes fraction'),
         horizontalalignment='left', verticalalignment='bottom')

# Set plot label and axis labels
ax.set(title='Number of Total Failures Observed')
ax.set(xlabel='Number of Failures', ylabel='Count')

# Save plot as .jpg ready for MS PowerPoint widescreen presentation
if save_plots == 'Y':
   plt.savefig('08 Number of Total Failures Observed - Case ' + str(case) + '.jpg', dpi=1200)
else:
   pass

# Display
plt.show()
```



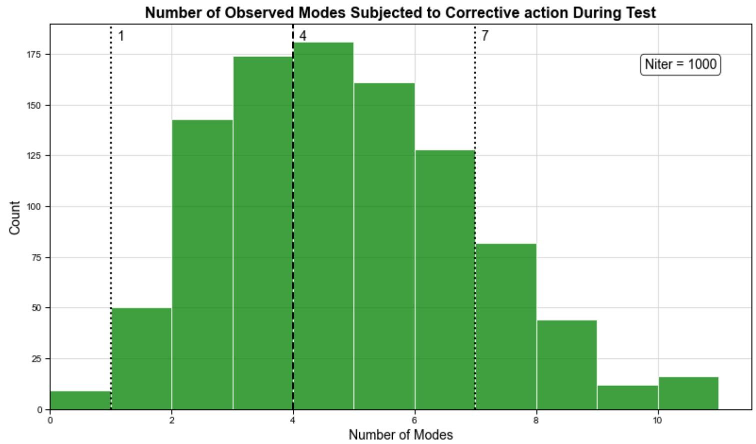

```
# Total Failures Observed Before Fixes Implementation Probability Density Distribution
# Set plot size
fig, ax = plt.subplots(figsize=(width, height))

plt.axvline(n1_lcb, lw=2, ls=':')
plt.axvline(n1_mean, lw=2, ls='--')
plt.axvline(n1_ucb, lw=2, ls=':')

sns.histplot(pp_n1, color='red', binwidth=1)

# Plot number of observed modes distribution
#sns.kdeplot(pp_n1,
#           linestyle='-',
#           linewidth=3.0,
#           bw_adjust=2,
#           color='red',
#           shade=True,
#           alpha=0.3,
#           label='Bayes Empirical Approach')

# Control x and y limits mode failure distribution plot
ax.set(ylim=(0.0, None))
ax.set(xlim=(np.min(pp_n1), round(n1_ucb * 2)))

# Annotate mode failure probability parameter
plt.annotate('Niter = ' + str(niter), xy=(0.95, 0.91), xycoords='axes fraction',
          horizontalalignment='right', verticalalignment='top', bbox=bbox_args)

# Annotate LCB, mean, UCB
plt.annotate(str(n1_lcb), xy=(n1_lcb + 0.1, 0.95), xycoords=('data','axes fraction'),
          horizontalalignment='left', verticalalignment='bottom')
plt.annotate(str(n1_mean), xy=(n1_mean + 0.1, 0.95), xycoords=('data','axes fraction'),
          horizontalalignment='left', verticalalignment='bottom')
plt.annotate(str(n1_ucb), xy=(n1_ucb + 0.1, 0.95), xycoords=('data','axes fraction'),
          horizontalalignment='left', verticalalignment='bottom')

# Set plot label and axis labels
ax.set(title='Number of Total Failures Observed Before Fixes Implementation')
ax.set(xlabel='Number of Failures', ylabel='Count')

# Save plot as .jpg ready for MS PowerPoint widescreen presentation
if save_plots == 'Y':
   plt.savefig('09 Number of Total Failures Observed Before Fixes Implementation - Case ' + str(case) + '.jpg', dpi=1200)
else:
   pass

# Display
plt.show()
```



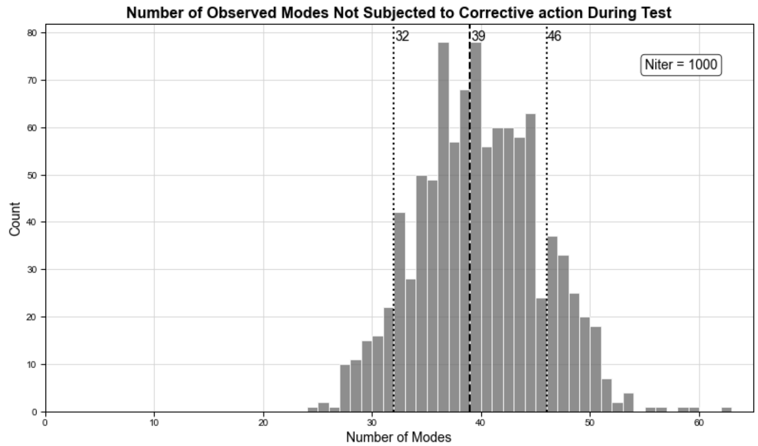

```
# Number of Total Failures Observed After Fixes Implementation
# Set plot size
fig, ax = plt.subplots(figsize=(width, height))

plt.axvline(n2_lcb, lw=2, ls=':')
plt.axvline(n2_mean, lw=2, ls='--')
plt.axvline(n2_ucb, lw=2, ls=':')

sns.histplot(pp_n2, color='violet', binwidth=1)

# Plot number of observed modes distribution
#sns.kdeplot(pp_n2,
#           linestyle='-',
#           linewidth=3.0,
#           bw_adjust=2,
#           color='violet',
#           shade=True,
#           alpha=0.3,
#           label='Bayes Empirical Approach')

# Control x and y limits mode failure distribution plot
ax.set(ylim=(0.0, None))
ax.set(xlim=(0.0, None))

# Annotate mode failure probability parameter
plt.annotate('Niter = ' + str(niter), xy=(0.95, 0.91), xycoords='axes fraction',
           horizontalalignment='right', verticalalignment='top', bbox=bbox_args)

# Annotate LCB, mean, UCB
#plt.annotate(str(n2_lcb), xy=(n2_lcb + 0.1, 0.95), xycoords=('data','axes fraction'),
#            horizontalalignment='left', verticalalignment='bottom')
plt.annotate(str(n2_mean), xy=(n2_mean + 0.1, 0.95), xycoords=('data','axes fraction'),
           horizontalalignment='left', verticalalignment='bottom')
plt.annotate(str(n2_ucb), xy=(n2_ucb + 0.1, 0.95), xycoords=('data','axes fraction'),
           horizontalalignment='left', verticalalignment='bottom')

# Set plot label and axis labels
ax.set(title='Number of Total Failures Observed After Fixes Implementation')
ax.set(xlabel='Number of Failures', ylabel='Count')

# Save plot as .jpg ready for MS PowerPoint widescreen presentation
if save_plots == 'Y':
   plt.savefig('10 Number of Total Failures Observed After Fixes Implementation - Case ' + str(case) + '.jpg', dpi=1200)
else:
   pass

# Display
plt.show()
```




```
# Number of Modes Subjected to Corrective Action During Test

# Set plot size
fig, ax = plt.subplots(figsize=(width, height))

plt.axvline(nfix_lcb, lw=2, ls=':')
plt.axvline(nfix_mean, lw=2, ls='--')
plt.axvline(nfix_ucb, lw=2, ls=':')

sns.histplot(pp_nfix, color='green', binwidth=1)

# Plot number of observed modes distribution
#sns.kdeplot(pp_nfix,
#          linestyle='-',
#          linewidth=3.0,
#          bw_adjust=2,
#          color='green',
#          shade=True,
#          alpha=0.3,
#          label='Bayes Empirical Approach')

# Control x and y limits mode failure distribution plot
ax.set(ylim=(0.0, None))
ax.set(xlim=(0.0, None))

# Annotate mode failure probability parameter
plt.annotate('Niter = ' + str(niter), xy=(0.95, 0.91), xycoords='axes fraction',
          horizontalalignment='right', verticalalignment='top', bbox=bbox_args)

# Annotate LCB, mean, UCB
plt.annotate(str(nfix_lcb), xy=(nfix_lcb + 0.1, 0.95), xycoords=('data','axes fraction'),
          horizontalalignment='left', verticalalignment='bottom')
plt.annotate(str(nfix_mean), xy=(nfix_mean + 0.1, 0.95), xycoords=('data','axes fraction'),
          horizontalalignment='left', verticalalignment='bottom')
plt.annotate(str(nfix_ucb), xy=(nfix_ucb + 0.1, 0.95), xycoords=('data','axes fraction'),
          horizontalalignment='left', verticalalignment='bottom')

# Set plot label and axis labels
ax.set(title='Number of Observed Modes Subjected to Corrective action During Test')
ax.set(xlabel='Number of Modes', ylabel='Count')

# Save plot as .jpg ready for MS PowerPoint widescreen presentation
if save_plots == 'Y':
   plt.savefig('11 Number of Observed Modes Subjected to Corrective action During Test - Case ' + str(case) + '.jpg', dpi=1200)
else:
   pass

# Display
plt.show()
```

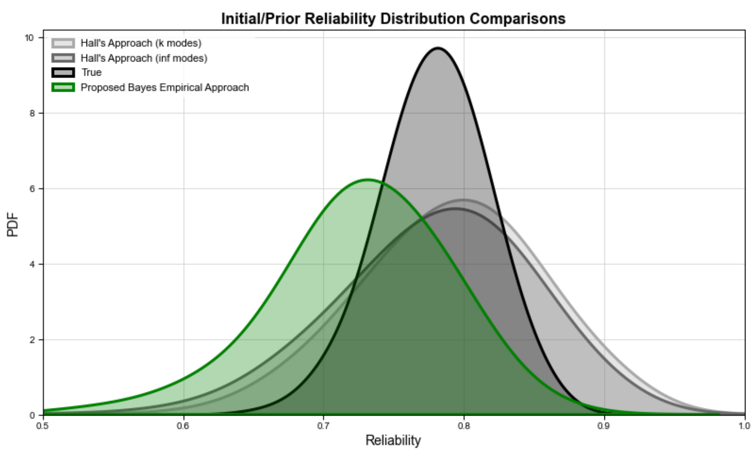



```
# Number of Modes Not Subjected to Corrective Action During Test
# Set plot size
fig, ax = plt.subplots(figsize=(width, height))

plt.axvline(nnofix_lcb, lw=2, ls=':')
plt.axvline(nnofix_mean, lw=2, ls='--')
plt.axvline(nnofix_ucb, lw=2, ls=':')

sns.histplot(pp_nnofix, color='dimgray', binwidth=1)

# Plot number of observed modes distribution
#sns.kdeplot(pp_nnofix,
#           linestyle='-',
#           linewidth=3.0,
#           bw_adjust=2,
#           color='dimgray',
#           shade=True,
#           alpha=0.3,
#           label='Bayes Empirical Approach')

# Control x and y limits mode failure distribution plot
ax.set(ylim=(0.0, None))
ax.set(xlim=(0.0, None))

# Annotate niter
plt.annotate('Niter = ' + str(niter), xy=(0.95, 0.91), xycoords='axes fraction',
           horizontalalignment='right', verticalalignment='top', bbox=bbox_args)

# Annotate LCB, mean, UCB
plt.annotate(str(nnofix_lcb), xy=(nnofix_lcb + 0.1, 0.95), xycoords=('data','axes fraction'),
           horizontalalignment='left', verticalalignment='bottom')
plt.annotate(str(nnofix_mean), xy=(nnofix_mean + 0.1, 0.95), xycoords=('data','axes fraction'),
           horizontalalignment='left', verticalalignment='bottom')
plt.annotate(str(nnofix_ucb), xy=(nnofix_ucb + 0.1, 0.95), xycoords=('data','axes fraction'),
           horizontalalignment='left', verticalalignment='bottom')

# Set plot label and axis labels
ax.set(title='Number of Observed Modes Not Subjected to Corrective action During Test')
ax.set(xlabel='Number of Modes', ylabel='Count')

# Save plot as .jpg ready for MS PowerPoint widescreen presentation
if save_plots == 'Y':
    plt.savefig('12 Number of Observed Modes Not Subjected to Corrective action During Test - Case ' + str(case) + '.jpg',
dpi=1200)
else:
    pass

# Display
plt.show()
```



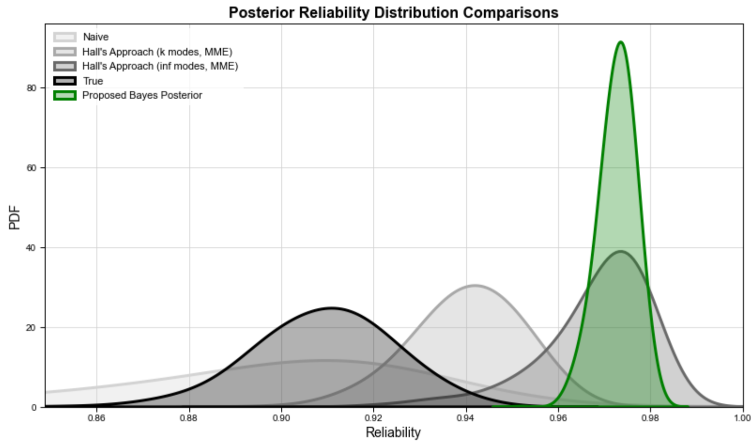

```
# Comparison of Initial Reliability Distributions
# Set plot size
fig, ax = plt.subplots(figsize=(width, height))

# Plot Hall's init reliability estimator distribution (k modes)
sns.kdeplot(pp_R_hall_k,
        linestyle='-',
        linewidth=3.0,
        bw_adjust=2,
        color='darkgrey',
        shade=True,
        alpha=0.3,
        label="Hall's Approach (k modes)")

# Plot Hall's init reliability estimator distribution (inf modes)
sns.kdeplot(pp_R_hall_inf,
        linestyle='-',
        linewidth=3.0,
        bw_adjust=2,
        color='dimgrey',
        shade=True,
        alpha=0.3,
        label="Hall's Approach (inf modes)")

# Plot true init reliability distribution
sns.kdeplot(pp_R_true,
        linestyle='-',
        linewidth=3.0,
        bw_adjust=2,
        color='black',
        shade=True,
        alpha=0.3,
        label='True')

# Plot proposed init reliability estimator distribution
sns.kdeplot(pp_R_emp,
        linestyle='-',
        linewidth=3.0,
        bw_adjust=2,
        color='green',
        shade=True,
        alpha=0.3,
        label='Proposed Bayes Empirical Approach')

# Control x and y limits
ax.set(ylim=(0, None))
ax.set(xlim=(0.5, 1))

# Send grid behind data plots
ax.grid(zorder=60)
```

```
# Set plot label
ax.set(title='Initial/Prior Reliability Distribution Comparisons')

# Set axes labels
ax.set(xlabel='Reliability', ylabel='PDF')

# Plot legend and format
plt.legend(loc='upper left', edgecolor='w', ncol=1)

# Save plot as .jpg ready for MS PowerPoint widescreen presentation
if save_plots == 'Y':
    plt.savefig('13 Initial_Prior Reliability Distribution Comparisons - Case ' + str(case) + '.jpg', dpi=1200)
else:
    pass

plt.show()
```



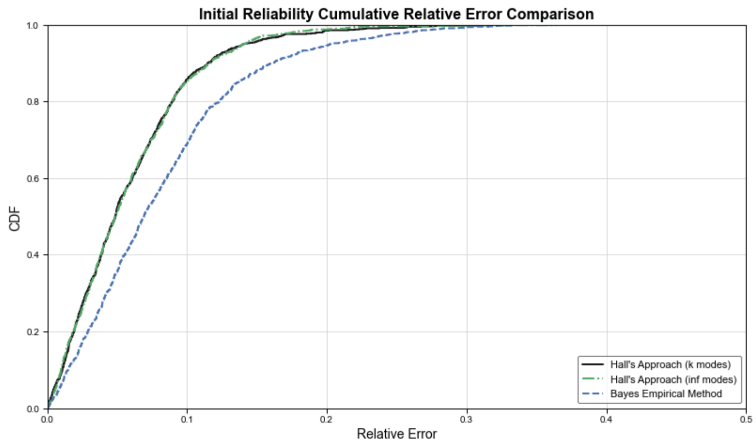

```
# Comparison of Posterior Reliability Distributions

# Set plot size


fig, ax = plt.subplots(figsize=(width, height))




# Plot naive reliability
sns.kdeplot(pp_naive,
        linestyle='-',
        linewidth=3.0,
        bw_adjust=2,
        color='lightgray',
        shade=True,
        alpha=0.3,
        label='Naive')

#sns.histplot(data=pp_post_hall_k)

# Plot Hall's k modes posterior reliability
sns.kdeplot(pp_post_hall_k,
                linestyle='-',
                linewidth=3.0,
                bw_adjust=2,
                color='darkgray',
                shade=True,
                alpha=0.3,
                label="Hall's Approach (k modes)")
```

```
# Plot Hall's inf modes posterior reliability
sns.kdeplot(pp_post_hall_inf,
            linestyle='-',
            linewidth=3.0,
            bw_adjust=2,
            color='dimgray',
            shade=True,
            alpha=0.3,
            label="Hall's Approach (inf modes)")

# Plot true init reliability distribution
sns.kdeplot(pp_true_post,
            linestyle='-',
            linewidth=3.0,
            bw_adjust=2,
            color='black',
            shade=True,
            alpha=0.3,
            label='True')

#pp_Bayes_mean=pp_Bayes_mean
sns.kdeplot(pp_Bayes_mean,
        linestyle='-',
        linewidth=3.0,
        bw_adjust=2,
        color='green',
        shade=True,
        alpha=0.3,
        label='Proposed Bayes Posterior')

# Control x and y limits
ax.set(ylim=(0, None))
ax.set(xlim=(np.min(pp_true_post), 1))

# Set plot label
ax.set(title='Posterior Reliability Distribution Comparisons')

# Set axes labels
ax.set(xlabel='Reliability', ylabel='PDF')

# Plot legend and format
plt.legend(loc='upper left', edgecolor='w', ncol=1)

# Save plot as .jpg ready for MS PowerPoint widescreen presentation
if save_plots == 'Y':
    plt.savefig('14 Posterior Reliability Distribution Comparisons - Case ' + str(case) + '.jpg', dpi=1200)
else:
    pass

plt.show()
```



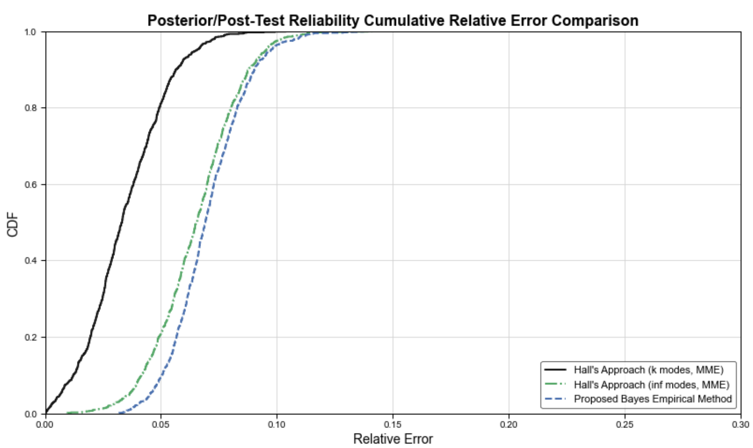

```
# Plot Relative Error between Prior Estimates and True Initial Reliability
# Set plot size
fig, ax = plt.subplots(figsize=(width, height))

# Initial reliability relative error using Hall k approach
hall_k_rel_error = np.empty(niter)
for i in range(niter):
    hall_k_rel_error[i] = abs(pp_R_hall_k[i] - pp_R_true[i]) / pp_R_true[i]

# Initial reliability relative error using Hall inf approach
hall_inf_rel_error = np.empty(niter)
for i in range(niter):
    hall_inf_rel_error[i] = abs(pp_R_hall_inf[i] - pp_R_true[i]) / pp_R_true[i]

# Initial reliability relative error using Bayes empirical method
Bayes_emp_rel_error = np.empty(niter)
for i in range(niter):
    Bayes_emp_rel_error[i] = abs(pp_R_emp[i] - pp_R_true[i]) / pp_R_true[i]

# Plot Hall k approach relative error
sns.ecdfplot(hall_k_rel_error,
            color='k',
            linestyle='-',
            linewidth=2.0,
            label="Hall's Approach (k modes)")

# Plot Hall inf approach relative error
sns.ecdfplot(hall_inf_rel_error,
            color='g',
            linestyle='-.',
            linewidth=2.0,
            label="Hall's Approach (inf modes)")

# Plot Bayes empirical method relative error
sns.ecdfplot(Bayes_emp_rel_error,
            color='b',
            linestyle='--',
            linewidth=2.0,
            label='Bayes Empirical Method')

# Control x and y limits
ax.set(ylim=(0, 1.0))
ax.set(xlim=(0, 0.5))

# Set plot label
ax.set(title='Initial Reliability Cumulative Absolute Relative Error Comparison')

# Plot legend and format
plt.legend(loc='lower right', edgecolor='k')
```

```
# Set axes labels
ax.set(xlabel='Relative Error', ylabel='CDF')

# Save plot as .jpg ready for MS PowerPoint widescreen presentation
if save_plots == 'Y':
    plt.savefig('15 Initial Reliability Cumulative Relative Error Comparison - Case ' + str(case) + '.jpg', dpi=1200)
else:
    pass

plt.show()
```



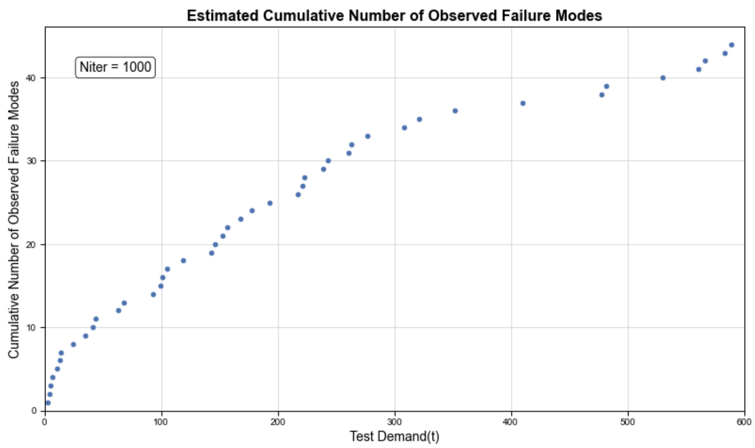

```
# Plot Relative Error between Posterior Estimates and True Posterior Reliability
# Set plot size
fig, ax = plt.subplots(figsize=(width, height))

# Post-test reliability relative error using naive approach
#naive_rel_error = np.empty(niter)
#for i in range(niter):
#    naive_rel_error[i] = abs(pp_naive[i] - pp_true_post[i]) / pp_true_post[i]

# Posterior reliability relative error using Hall k approach
hall_k_rel_error = np.empty(niter)
for i in range(niter):
    hall_k_rel_error[i] = abs(pp_post_hall_k[i] - pp_true_post[i]) / pp_true_post[i]

# Posterior reliability relative error using Hall inf approach
hall_inf_rel_error = np.empty(niter)
for i in range(niter):
    hall_inf_rel_error[i] = abs(pp_post_hall_inf[i] - pp_true_post[i]) / pp_true_post[i]

# Posterior reliability relative error using Bayesian method
Bayes_emp_rel_error = np.empty(niter)
for i in range(niter):
    Bayes_emp_rel_error[i] = abs(pp_Bayes_mean[i] - pp_true_post[i]) / pp_true_post[i]

# Plot naive approach relative error
#sns.ecdfplot(naive_rel_error,
#            color='darkgray',
#            linestyle=':',
#            linewidth=2.0,
#            label='Naive Approach')

# Plot Hall k approach relative error
sns.ecdfplot(hall_k_rel_error,
            color='k',
            linestyle='-',
            linewidth=2.0,
            label="Hall's Approach (k modes)")

# Plot Hall inf approach relative error
```

```
sns.ecdfplot(hall_inf_rel_error,
        color='g',
        linestyle='-.',
        linewidth=2.0,
        label="Hall's Approach (inf modes)")

# Plot Bayes empirical method relative error
sns.ecdfplot(Bayes_emp_rel_error,
        color='b',
        linestyle='--',
        linewidth=2.0,
        label='Proposed Bayes Approach')

# Control x and y limits
ax.set(ylim=(0, 1.0))
ax.set(xlim=(0, 0.3))

# Set plot label
ax.set(title='Posterior/Post-Test Reliability Cumulative Absolute Relative Error Comparison')

# Plot legend and format
plt.legend(loc='lower right', edgecolor='k')

# Set axes labels
ax.set(xlabel='Relative Error', ylabel='CDF')

# Save plot as .jpg ready for MS PowerPoint widescreen presentation
if save_plots == 'Y':
  plt.savefig('16 Posterior_Post Test Reliability Cumulative Relative Error Comparison - Case ' + str(case) + '.jpg', dpi=1200)
else:
  pass

plt.show()
```




```
# Plot Expected Number of Correctable Failure Modes (From Random Sample)

# Set plot size
fig, ax = plt.subplots(figsize=(width, height))

y = np.linspace(1, len(rand_sims[0]), len(rand_sims[0]))
plt.scatter(rand_sims, y)

# Control x and y limits mode failure distribution plot
ax.set(ylim=(0.0, None))
ax.set(xlim=(0.0, T))

# Annotate niter
plt.annotate('Niter = ' + str(niter), xy=(0.05, 0.91), xycoords='axes fraction',
        horizontalalignment='left', verticalalignment='top', bbox=bbox_args)
```

```
# Set plot label and axis labels
ax.set(title='Estimated Cumulative Number of Observed Failure Modes')
ax.set(xlabel='Test Demand(t)', ylabel='Cumulative Number of Observed Failure Modes')

# Save plot as .jpg ready for MS PowerPoint widescreen presentation
if save_plots == 'Y':
    plt.savefig('17 Cumulative Number of Observed Failure Modes versus Expected Number of Correctable Modes - Case ' +
str(case) + '.jpg', dpi=1200)
else:
    pass

# Display
plt.show()
```



Bibliography

[1] E. Zio, "Reliability engineering: Old problems and new challenges," *Reliability Engineering & System Safety,* vol. 94, no. 2, pp. 125-141, 2009.

[2] HBM Prenscia, "Reliability Growth Analysis," HBM Prenscia, n.d.. [Online]. Available: https://www.weibull.com/basics/growth.htm. [Accessed 22 October 2020].

[3] HBM Prenscia, "RGA Overview," HBM Prenscia, 11 December 2015. [Online]. Available: http://reliawiki.org/index.php/RGA_Overview. [Accessed 22 October 2020].

[4] ReliaSoft Corporation, Reliability Growth and Repairable System Analysis Reference, Tucson, AZ: ReliaSoft Corporation, 2019.

[5] US Dept of Defense, Military Handbook 189: Reliability Growth Management, Aberdeen Proving Ground, MD: US Dept of Defense, 1981.

[6] US Dept of Defence, Military Handbook 189A: Reliability Growth Management (MIL-HDBK-189A), Aberdeen Provding Ground, MD: US Dept of Defense, 2009.

[7] US Dept of Defense, Military Handbook 189C: Reliability Growth Management (MIL-HDBK-189C), Aberdeen Proving Ground, MD: US Dept of Defense, 2011.

[8] M. Mobin and M. Hijawi, "A simulation-based approach to incorporate uncertainty in reliability growth planning (RGP)," in *WCX SAE World Congress Experience*, Virtual Conference, 2020.

[9] J. B. Hall, "Reliability growth planning for discrete-use systems," in *Proceedings of the Annual Reliability and Maintainability Symposium*, Lake Buena Vista, FL, 2011.

[10] R. Strunz and J. W. Herrmann, "Planning, tracking and projecting reliability growth: A Bayesian approach," in *Proceedings of the Annual Reliability and Maintainability Symposium*, Reno, NV, 2012.

[11] J. B. Hall, "Doctoral Dissertation: Methodology for evaluating reliability growth programs of discrete systems," University of Maryland, College Park, MD, 2008.

[12] M. Wayne, “Doctoral Dissertation: Methodology for assessing reliability growth using multiple information sources,” University of Maryland, College Park, MD, 2013.

[13] J. T. Duane, “Learning approach to reliability monitoring,” *IEEE Transactions on Aerospace,* vol. 2, no. 2, pp. 563-566, April 1964.

[14] J. D. Selby and S. G. Miller, “Reliability Planning and Management,” in *ASQC/SRE Seminar*, Niagara Falls, NY, 1970.

[15] L. Crow, “Reliability Analysis for Complex, Repairable Systems,” in *Reliability and Biometry*, Philadelphia, PA, Society for Industrial and Applied Mathematics, 1975, pp. 379-410.

[16] M. McCarthy, “Technical Report TR-555: Developing a Subsystem Reliability Growth Program Using the Subsystem Reliability Growth Planning Model,” US Dept of the Army, Aberdeen Proving Ground, MD, 1994.

[17] P. M. Ellner and J. B. Hall, “An approach to reliability growth planning based on failure mode discovery and correction using AMSAA projection methodology,” in *Proceedings of the Annual Reliability and Maintainability Symposium*, Newport Beach, CA, 2006.

[18] L. H. Crow, “Planning a reliability growth program utilizing historical data,” in *Proceedings of the Annual Reliability and Maintainability Symposium*, Lake Buena Vista, FL, 2011.

[19] M. Wayne, “Methodology for assessing reliability growth using multiple information sources,” Dr. Martin Wayne PhD Dissertation, University of Maryland, College Park, MD, 2013.

[20] M. Wayne, “Modeling uncertainty in reliability growth plans,” in *Proceedings of the Annual Reliability and Maintainability Symposium*, Reno, NV, 2018.

[21] W. J. Corcoran, H. Weingarten and P. W. Zehna, “Estimating reliability after corrective action,” *Management Science,* vol. 10, no. 4, pp. 601-813, 1964.

[22] L. H. Crow, “Technical Report TR-357: An Improved Methodology for Reliability Growth Projection,” U.S. Dept of the Army, Aberdeen Proving Ground, MD, 1982.

[23] P. M. Ellner and L. C. Wald, “AMSAA maturity projection model,” in *Proceedings of the Annual Reliability and Maintainability Symposium*, Washington, DC, 1995.

[24] J. A. Clark, “Modeling reliability growth late in development,” in *Proceedings of the Annual Reliability and Maintainability Symposium*, Washington, DC, 1999.

[25] P. M. Ellner and J. B. Hall, “Technical Report TR-751: AMSAA Maturity Projection Model Based on Stein Estimation,” U.S. Dept of the Army, Aberdeen Proving Ground, MD, 2004.

[26] C. M. Stein, “Estimation of the Mean of a Multivariate Normal Distribution,” *The Annals of Statistics,* vol. 9, no. 6, pp. 1135-1151, 1981.

[27] L. H. Crow, “An extended reliability growth model for managing and assessing corrective actions,” in *Proceedings of the Annual Reliability and Maintainability Symposium*, Los Angeles, CA, 2004.

[28] L. H. Crow, “Interim Note R-30: On Reliability Growth Tracking,” U.S. Dept of the Army, Aberdeen Proving Ground, MD, 1974.

[29] M. Wayne and P. M. Ellner, “Comparing the robustness of reliability growth projection models,” in *Proceedings of the Annual Reliability and Maintainability Symposium*, San Jose, CA, 2010.

[30] J. B. Hall and A. Mosleh, “A reliability growth projection model for one-shot systems,” *IEEE Transactions on Reliability,* vol. 57, no. 1, pp. 174-181, 2008.

[31] J. B. Hall and A. Mosleh, “An analytical framework for reliability growth of one-shot systems,” *Reliability Engineering and System Safety,* vol. 93, no. 11, pp. 1751-1760, 2008.

[32] J. B. Hall and A. Mosleh, “Bayesian methods for evaluating discrete reliability growth,” in *Proceedings of the Annual Reliability and Maintainability Symposium*, Fort Worth, TX, 2009.

[33] J. B. Hall, P. M. Ellner and A. Mosleh, “Reliability growth management metrics and statistical methods for discrete-use systems,” *Technometrics,* vol. 52, no. 4, pp. 379-389, 2010.

[34] ReliaSoft Corporation, “Reliability Growth & Repairable System Analysis Reference, Chapter 1: RGA Overview,” ReliaSoft Publishing, Tucson, AZ, 2015.

[35] J. B. Hall, P. M. Ellner and A. Mosleh, “Reliability growth management metrics and statistical methods for discrete-use systems,” *Technometrics,* vol. 52, no. 4, pp. 379-389, 2010.

[36] J. B. Hall, “Reliability growth planning for discrete-use systems,” in *Proceedings of the Annual Reliability and Maintainability Symposium*, Lake Buena Vista, FL, 2011.

[37] Quanterion Solutions Incorporated, “Models Commonly Used to Measure Reliability Growth,” Quanterion Solutions Incorporated, 26 April 2013. [Online]. Available: https://www.quanterion.com/models-commonly-used-to-measure-reliability-growth/#:~:text=The%20%E2%80%9Cfix%20effectiveness%20factor%E2%80%9D%20or,a%20completely%20ineffective%20corrective%20action.. [Accessed 7 April 2021].

[38] M. D. Springer, The Algebra of Random Variables, New York: Wiley & Sons, 1979.

[39] B. D. Carter and M. D. Springer, “The distribution of products, quotients and powers of independent H-function variates,” *Society of Industrial and Applied Mathematics (SIAM) Journal of Applied Mathematics,* vol. 33, no. 4, p. 542–558, 1977.

[40] L. H. Crow, “Estimating the management strategy for reliability growth planning,” in *Proceedings of the Annual Reliability and Maintainability Symposium*, Colorado Springs, CO, 2014.

[41] E. Jaynes, “Information theory and statistical mechanics,” *Physical Review,* vol. 106, no. 4, pp. 620-630, 1957.

[42] E. Jaynes, “Information theory and statistical mechanics II,” *Physical Review,* vol. 108, no. 2, pp. 171-190, 1957.

[43] S. Brown, “Doctoral Dissertation: Development and validation of methodology for fix effectiveness projection during product development,” University of Maryland, College Park, MD, 2009.

[44] U.S. Dept of Defense, Documentation of verification, Validation, and Accreditation (VV&A) for Models and Simulations, Alexandria, VA: U.S. Dept of Defense, 2012.

[45] R. Cox, "Probability, frequency, and reasonable expectation," *American Journal of Physics,* vol. 14, no. 1, pp. 1-10, 1946.

[46] E. Jaynes, "Bayesian methods: General background," in *Maximum-Entropy and Bayesian Methods in Applied Statistics*, Cambridge, UK, Cambridge University Press, 1986.

[47] B. de Finetti, Theory of Probability: A Critical Introductory Treatment, Chichester, UK: John Wiley & Sons, 2017.

[48] J. Neyman, "Outline of a theory of statistical estimation based on the classical theory of probability," *Philosophical Transactions of the Royal Society,* vol. 236, no. 767, pp. 333-380, 1937.

[49] J. Jeon and S. Ahn, "Bayesian methods for reliability demonstration test for finite population using lot and sequential sampling," *Sustainability,* vol. 10, no. 10, p. 3671 (Article Number), 2018.

[50] H. Mignon, G. Dani and L. Francisco, Statistical Inference: An Integrated Approach, London, UK: CRC Press, 2014.

[51] C. Bishop, Neural Networks for Pattern Recognition, New York, USA: Oxford University Press, 1995.

[52] J. Hall, M. Ellner and A. Mosleh, "Reliability growth management metrics and statistical methods for discrete-use systems," *Technometrics,* vol. 52, no. 4, pp. 379-389, 2010.

[53] M. Wayne and M. Modarres, "A Bayesian model for complex system reliability growth under arbitrary corrective actions," *IEEE Transactions on Reliability,* vol. 64, no. 1, pp. 206-220, 2015.

[54] C. Cormick, The Science of Communicating Science: The Ultimate Guide, Melbourne, Australia: CSIRO Publishing, 2019.

[55] D. R. Laker and J. L. Powell, "The differences between hard and soft skills and their relative impact on training transfer," *Human Resource Development Quarterly,* vol. 22, no. 1, pp. 111-122, 2011.

[56] N. Luhmann, “What is communication?,” *Communication Theory,* vol. 2, no. 3, pp. 251-259, 1992.

[57] The Kautilya Society, “The Difference between "Communication" and "Information",” Kautilya Society of Intercultural Dialogue, [Online]. Available: http://www.kautilyasociety.com/tvph/varieties/coomunication_information.htm#:~:text=Communication%20is%20different%20from%20information%3A%20communication%20is%20active,%27information%3A%20it%20is%20the%20transmission%20of%20a%20message. [Accessed 2021 July 18].

[58] R. D. Aines and A. L. Aines, Championing Science: Communicating Your Ideas to Decision Makers, Oakland, CA: University of California Press, 2019.

[59] C. Atkin and V. Freimuth, “Formative evaluation research in campaign design,” in *Public Communication Campaigns*, Thousand Oaks, CA, SAGE Publishing, 2001, pp. 125-145.

[60] B. Dervin and M. Frenette, “Applying sense-making methodology: Communicating communicatively with audiences as listeners, learners, teachers, confidantes,” in *Public Communication Campaigns*, Thousand Oaks, CA, SAGE Publishing, 2001, pp. 69-87.

[61] J. N. Kim, H. Shen and S. E. Morgan, “Information behaviors and problem chain recognition effect: Applying situational theory of problem solving in organ donation issues,” *Health Communication,* vol. 26, no. 2, pp. 171-184, 2011.

[62] Business School, University of Navarra, “The 10 Most Common Mistakes in Decision-Making,” University of Navarra, Barcelona, 2017.

[63] A. Huertas, “Developing effective messages in science communication,” medium.com, 2016.

[64] COMPASS Science Communication, Inc., “The Message Box,” COMPASS Science Communication, Inc., [Online]. Available: https://www.compassscicomm.org/leadership-development/the-message-box/#:~:text=Every%20scientist%20we%20work%20with%20starts%20with%20a,and%20why%20it%20matters%20for%20a%20particular%20audience.. [Accessed 20 July 2021].

[65] K. Fog, C. Budtz, P. Munch and S. Blanchette, Storytelling: Branding in Practice, Copenhagen, Denmark: Sigma, 2010.

[66] J. Downs, “Prescriptive scientific narratives for communicating usable science,” *Proceedings of the National Academy of Sciences of the United States of America,* vol. 111, no. Supplement 4, pp. 13627-33, 2014.

[67] T. Greenhalgh, “Storytelling should be targeted where it is known to have greatest added value,” *Med Ed,* vol. 35, no. 9, pp. 818-819, 2001.

[68] C. Rapp, “Aristotle's Rhetoric,” in *Stanford Encylopedia of Philosophy*, 2010.

[69] M. Dahlstrom, “Using narratives and storytelling to communicate science with non-expert audiences,” *Proceedings of the National Academy of Sciences of the United States of America,* vol. 111, no. Supplement 4, pp. 13614-20, 2014.

[70] J. Druckman and T. Bolsen, “Framing, motivated reasoning, and opinions about emerging technologies,” *J Comm,* vol. 61, no. 4, pp. 659-688, 2011.

[71] US Dept of Defense, Military Handbook 189C: Reliability Growth Management (MIL-HDBK-189C), Aberdeen Proving Ground, Maryland, USA: US Dept of Defense, 2011.

[72] US Dept of Army, “Guidance for Assessing the Likelihood that a System will Demonstrate its Reliability Requirement during Initial Operational Test,” US Dept of Army, n.d..

[73] K. Ramachandran and C. Tsokos, Mathematical Statistics with Applications in R, Academic Press, 2015.

[74] D. Navarro, “Assessing the Reliability of a Measurement,” LibreTexts, 18 August 2020. [Online]. Available: https://stats.libretexts.org/@go/page/3946. [Accessed 23 July 2021].

[75] J. Shin and D. Zurakowski, “Null hypotheses, interval estimation, and Bayesian analysis,” *Otolaryngol Head Neck Surg.,* vol. 157, no. 6, pp. 919-920, 2017.

[76] R. Brightwell and R. Dransfield, Avoiding and Detecting Statistical Malpractice: Data and Analysis for Biologists, with R, InfluentialPoints, 2013.

[77] J. Neyman, “Outline of a theory of statistical estimation based on the classical theory of probability,” *Philo Trans Royal Soc London, Series A, Mathematical and Physical Sciences,* vol. 236, no. 767, pp. 333-380, 1937.

[78] Y. Dodge, The Oxford Dictionary of Statistical Terms, Oxford University press, 2006.

[79] M. Glaser, T. Langer and M. Webber, "True overconfidence in interval estimates: Evidence based on a new measure of miscalibration," *J Behav Decision Making,* vol. 26, no. 5, pp. 405-417, 2013.

[80] S. Greenland, "Interval estimation by simulation as an alternative to and extension of confidence intervals," *Intl J Epidem,* vol. 33, no. 6, pp. 1389-1397, 2004.

[81] J. Steiger and R. Fouladi, "Noncentrality Interval Estimation and The Evaluation of Statistical Models," in *What if There Were No Significance Tests?*, Taylor and Francis, 2016, pp. 197-230.

[82] A. Chaturvedi and A. Malhotra, "Estimation and testing procedures for the reliability functions of a family of lifetime distributions based on records," *Int J Syst Assur Eng Manag,* vol. 8, pp. 836-848, 2017.